

[illegible][illegible]

1950 1960 1970 1980 1990 2000 2010
CAGAAATGCAAGATGAAGTCAAAACTGTTGGTGAGGAAACAAAAGCCAGAGAAACCAAGCGTAGGCTGGGATA
1920 1930 1940 1950 1960 1970 1980
CAGAAATGCAAGATGAAGTCAAAACTGTTGGTGAGGAAACAAAAGCCAGAGAAACCAAGCGTAGGCTGGGATA
1990 2000 2010 2020 2030 2040 2050
CTTCAGTGTCTGGGAAGCACTGATTTGTTCGGATCATCCAAAGAGAGAGCAAGGAAGCATCTCTTCAG
2020 2030 2040 2050 2060 2070 2080
CTTCAGTGTCTGGGAAGCACTGATTTGTTCGGATCATCCAAAGAGAGAGCAAGGAAGCATCTCTTCAG
2090 2100 2110 2120 2130 2140 2150 2160
ATGATCAAGAGAGGCCAAGGACACTGGAGGGGACAGTCAACAGAGAGAGAGGCGGACGACAAAGCAAGAAAG
2170 2180 2190 2200 2210 2220 2230
CGGAAACAGACGCTGTTCTTCAGCACCCAGAGAGGACCAAGCGCAAGGAAGTTCTCTCACCCGAGCCAG
2130 2140 2150 2160 2170 2180 2190 2200
CGGAAACAGACGCTGTTCTTCAGCACCCAGAGAGGACCAAGCGCAAGGAAGTTCTCTCACCCGAGCCAG
2210 2220 2230 2240 2250 2260 2270
CGGAAAGCCCTTCCGAAGGGGAAGGTGTTCCACTTGGGAGTCAATTTAAAGATTAGTCACTCCAAAGAAAA
2280 2290 2300 2310 2320 2330 2340
AATCCAAAGTCAAAACTGGAAGAGAAAGCCGAAGACTCTAGTGTAGAGCAGTTGTCTCCACTGAGATCCAAACCGA
2350 2360 2370 2380 2390 2400 2410
GTAGAGAAGAAATCTTGGGTTTCATTTAAGAAATTCATCCCGGACCGCGAAGAAAGGCGAGACGGGAAGC
2420 2430 2440 2450 2460 2470 2480
GTAGAGAAGAAATCTTGGGTTTCATTTAAGAAATTCATCCCGGACCGCGAAGAAAGGCGAGACGGGAAGC
2490 2500 2510 2520 2530 2540 2550 2560
AAGAAACAGCCACTGTGGAAGACTCAGGCGCAGTGGAGATAAATGAGGACGACCTTAATGTCTCCACCGCTCG
2570 2580 2590 2600 2610 2620 2630
TGCTGTGTCTGAGTATAATGCAAGTGGAGGGGAGAGATGGAAGCCAGGGGAATACGGAGTGCCCCAGC
2640 2650 2660 2670 2680 2690 2700
TGCTGTGTCTGAGTATAATGCAAGTGGAGGGGAGAGATGGAAGCCAGGGGAATACGGAGTGCCCCAGC
2710 2720 2730 2740 2750 2760 2770
TGCTGTGTCTGAGTATAATGCAAGTGGAGGGGAGAGATGGAAGCCAGGGGAATACGGAGTGCCCCAGC
2780 2790 2800
TTGAACACAGCGGGAGAGCCATGCCACTCTTGAAGAGGTCACTGAAAGAGATCATTTGCAAGAAAA
2810 2820 2830 2840 2850 2860 2870 2880
TTGAACACAGCGGGAGAGCCATGCCACTCTTGAAGAGGTCACTGAAAGAGATCATTTGCAAGAAAA

2810 2820 2830 2840 2850 2860 2870 2880
CTCCTGTCTCACCCAGAGCTTACAGAGGGTAAAGATGCCCATGACGACATGGTCACCAAGTGAAGTGATT
2890 2900 2910 2920 2930 2940 2950
TCACCTCAGAAAGCTGTGACAGCACAGAGACCTCAGAGGCTTCGGTACTGAAAGATTACCGAAGCATCGG
2960 2970 2980 2990 3000 3010 3020
GGGCCAAGAGAGACCAAGAGCATGGTCCGAGTTTCCAGCTGACTGATCCCCAGACACCAAGAGAAAG
3030 3040 3050 3060 3070 3080 3090
CCACCCAGTTCAGAGAGTAGAGAGTGGTGTCTAGATACAGAAGAGAGGCGCCAGACGAGGCCATCC
3100 3110 3120 3130 3140 3150 3160
TCCAAAGCCCTTCAGACCAAGGTGAAGAGGAGTCCAGGTGCTTCAACCCAGACTGTGCAGAGAACGGGGT
3170 3180 3190 3200 3210 3220 3230 3240
CAAAAGCACTGAGAGAGGTTGAGAGGAGTAGAGGAGGACTCCGAAGTGTGGTTCGGAGAAAGAGAGACG
3250 3260 3270 3280 3290 3300 3310
TTATGCGGAAAGGCCCTGCGAGGAGTGGAGCTGAGCATCTTGACAGGGCTCTGAGACTGGACAGGCTA
3320 3330 3340 3350 3360 3370 3380
CTCCAGAGAGCCTTGAAGTTCCTGAAGTCAACGCGAGATGTAGACCATGTGCGCAGTGCAGGTATCAAGC
3390 3400 3410 3420 3430 3440 3450
TCCAGCAGTGTATGAAACAGCGCCGTGGCCCTTGAGTCAATCCGAAACCTTTGACAGAGTGAACAATGGAA
3460 3470 3480 3490 3500 3510 3520
GCACTCCCTTACAGATTACAGACTGAGAGTGGGACACAGCAAGATGAAACCTTTGACAGCAGGACAGTA
3530 3540 3550 3560 3570 3580 3590 3600
AAGCCACTGACCTGTGAGGAGTCAAGGTCACAGAGAGAGGCGGCTACTGCTCAGAAAGAGAGGCTT
3610 3620 3630 3640 3650 3660 3670
CGACACTACTAATATGTTCCAGCCAGAGAAACATGGGGAAGAACCCAGGAAGAGATGTTCTTGAACCTA
3680 3690 3700 3710 3720 3730 3740
CGACACTACTAATATGTTCCAGCCAGAGAAACATGGGGAAGAACCCAGGAAGAGATGTTCTTGAACCTA

CAACAGCAAGAGCTTATCTGCTGAGCGCCGCCCCCTTCTGSCAAAGACTAGAGTGGATCAAGAGGGTGTAGGTTG
3750 3760 3770 3780 3790 3800 3810
ATTGGTTGGAATGGAGAAAAAGTCAAAAGAAACAGAGAGTGTGTTTATACATCTGGACCACTCACTCAAAAGG
3820 3830 3840 3850 3860 3870 3880
CTGCTGATGTGACATATGACATGACATGAAGTGAATGGAGGTGGCCGGGGGTGTAGAGAAAAAGAGATATCTGAAGTGC
3890 3900 3910 3920 3930 3940 3950
AAGAGTTTATGACCTTGGAGGAGGAGAGATGGAAATCTGAGCTTGAAGAAAGAGAGAGGAGCAAAAGCCAGAGC
3960 3970 3980 3990 4000 4010 4020 4030
AGAGCTTTAGCTTGGAGGAGGAGAGATGAATGAACCTTGAAGAAAGAAAGAGAGAGCAAAAGCCAGAGC
3930 3940 3950 3960 3970 3980 3990 4000
AAGTGTGATGAAGAAAGGTGTAGACGAAACAGCCGCTCTGAGGATGAAGAACTTACGGAGAGCCAGTCTTGA
4040 4050 4060 4070 4080 4090 4100
CAATTGACATGCCCACTCAGAGAGGGGAGAAAGCACTGGAGAGCCCTTGAGAGAAACCTTCTCTCCAGACC
4110 4120 4130 4140 4150 4160 4170
AAGACAAAGCAGAGTTCATAGAGATTCAGAGTTCAGAAAGCCTTGACACACACATGACTTCAACAGAGAAAGCTG
4180 4190 4200 4210 4220 4230 4240
TGGAAAAGTTCATAGAAAGGTTGTATTTTACAGACACAGTGAAGTCCAGAGTGTGATGAGCAACTTATAT
4250 4260 4270 4280 4290 4300 4310 4320
TACACAGTGAAGTCTCTCTGCAACGGGTGGCCACTGAGACTCTTTCAGATGTGAGAGACACGGTATCCCTCTGG
4330 4340 4350 4360 4370 4380 4390
GGCCTGAGTTCACGAGATTCATCTCCATCTATGTAATCTCTGCTCTCTGAAGACCTTCATCATCTGACCC
4400 4410 4420 4430 4440 4450 4460
TACAGAGAGAAATTAAGCGATCCCAAGAGAGCGATCTAAGAGAGAGGACAAAGCCAGAGTGTGTTCTGATGTG
4470 4480 4490 4500 4510 4520 4530
CTTACGCGCAAGAGAGAGTACAGCATTCGAAAAAGTCTCTCAAGGTGAACTTGAGATCTTGGAATTTGAGAGTAT
4540 4550 4560 4570 4580 4590 4600
AGGCAACAAGATTGTGCTGAACGTCATTCAGACACCCGTTGACAGTTTGACGTCATCAAGAAACAGCCCCCGG
4610 4620 4630 4640 4650 4660 4670
CAACAGCAAGAGCTTATCTGCTGAGCGCCGCCCCCTTCTGSCAAAGACTAGAGTGGATCAAGAGGGTGTAGGTTG
3650 3660 3670 3680 3690 3700 3710
CAACAGCAAGAGCTTATCTGCTGAGCGCCGCCCCCTTCTGSCAAAGACTAGAGTGGATCAAGAGGGTGTAGGTTG
3720 3730 3740 3750 3760 3770 3780
ACTGCTGATGATGAGAAAAAGTCAAAAGAAACAGAGAGTGTGTTTATACATCTGGACCACTCACTCAAAAGG
3790 3800 3810 3820 3830 3840 3850
CTGCTGATGTGACATATGACATGACATGAAGTGAAGTGGAGTGGCCGGGTGTAGAGAAAGAGATATCTGAAGTGC
3860 3870 3880 3890 3900 3910 3920
AAGTGTGATGAAGAAAGGTGTAGACGAAACAGCCGCTCTGAGGATGAAGAACTTACGGAGAGCCAGTCTTGA
3930 3940 3950 3960 3970 3980 3990 4000
AAGTGTGATGAAGAAAGGTGTAGACGAAACAGCCGCTCTGAGGATGAAGAACTTACGGAGAGCCAGTCTTGA
4010 4020 4030 4040 4050 4060 4070
CACTTACATGCCCACTCAGAGAGGGGAGAAAGCACTGGAGAGCCCTTCTCTCCAGACC
4080 4090 4100 4110 4120 4130 4140
AAGACAAAGCAGAGTTCATAGAGATTCAGAGTTCAGAAAGCCTTGACACACACATGACTTCAACAGAGAAAGCTG
4150 4160 4170 4180 4190 4200 4210
TGGAAAAGTTCATAGAAAGGTTGTATTTTACAGACACAGTGAAGTCCAGAGTGTGATGAGCAACTTATAT
4220 4230 4240 4250 4260 4270 4280
TACACAGTGAAGTCTCTCTGCAACGGGTGGCCACTGAGACTCTTTCAGATGTGAGAGACACGGTATCCCTCTGG
4290 4300 4310 4320 4330 4340 4350 4360
GGCCTGAGTTCACGAGATTCATCTCCATCTATGTAATCTCTGCTCTCTGAAGACCTTCATCATCTGACCC
4370 4380 4390 4400 4410 4420 4430
TACAGAGAGAAATTAAGCGATCCCAAGAGAGCGATCTAAGAGAGAGGACAAAGCCAGAGTGTGTTCTGATGTG
4440 4450 4460 4470 4480 4490 4500
CTTACGCGCAAGAGAGTACAGCATTCGAAAAAGTCTCTCAAGGTGAACTTGAGATCTTGGAATTTGAGAGTAT
4510 4520 4530 4540 4550 4560 4570
AGGCAACAAGATTGTGCTGAACGTCATTCAGACACCCGTTGACAGTTTGACGTCATCAAGAAACAGCCCCCGG

AGGGCAACAAGTTGTGTGGAACGGCATTTGAGAAGCGTTGAACGATTTGCGACGTTACAGAAACAGCCCCG
4510 4520 4530 4540 4550 4560 4570
4610 4620 4630 4640 4650 4660 4670 4680
AAACTCATGCTTATGATTCACAGACCAGAGTTCCTGCATGACAGGCTTGAACAGAGGAGGCCCAACAGATGCT
AAAACATGCTTATGATTCACAGACCAGAGTTCCTGCATGACAGGCTTGAACAGAGGAGGCCCAACAGATGCT
4580 4590 4600 4610 4620 4630 4640
4690 4700 4710 4720 4730 4740 4750
GGACAAAAATGAAAGATCCCAAGATGAAACCCAGTGCAGGACCCAGAGAGGACTTGAAGTCTCGACCG
GGACAAAAATGAAAGATCCCAAGATGAAACCCAGTGCAGGACCCAGAGAGGACTTGAAGTCTCGACCG
4650 4660 4670 4680 4690 4700 4710 4720
4760 4770 4780 4790 4800 4810 4820
TTTCGAGAGCATGGGCTCAGGCTCGGAATGCTTTCGCGCTTCAGTTGAAACCGCGGTGTCAAAAGTAG
TTTCGAGAGCATGGGCTCAGGCTCGGAATGCTTTCGCGCTTCAGTTGAAACCGCGGTGTCAAAAGTAG
4730 4740 4750 4760 4770 4780 4790
4830 4840 4850 4860 4870 4880 4890
CATTTGAAGTCTGCTCTCTCAACCCAAAGATCCAAAGAGCAATGCTGTGATGCGCTCAGCTCCAAAGCT
CATTTGAAGTCTGCTCTCTCAACCCAAAGATCCAAAGAGCAATGCTGTGATGCGCTCAGCTCCAAAGCT
4800 4810 4820 4830 4840 4850 4860
4900 4910 4920 4930 4940 4950 4960
TACCCAGAGCAAGAGCCAGTGGCTCTGTGAAACCTTAAACCAAGAAATCCCAAGATCCCAAGCAAGCAAGCTTA
TACCCAGAGCAAGAGCCAGTGGCTCTGTGAAACCTTAAACCAAGAAATCCCAAGATCCCAAGCAAGCAAGCTTA
4870 4880 4890 4900 4910 4920 4930
4970 4980 4990 5000 5010 5020 5030 5040
CCAGAGAGGCGCATCCCCCAAAAGTTGAGTGTCCAGAAAGAAATGTTACCAAGTCAAGTCAAAAGAGACA
CCAGAGAGGCGCATCCCCCAAAAGTTGAGTGTCCAGAAAGAAATGTTACCAAGTCAAGTCAAAAGAGACA
4940 4950 4960 4970 4980 4990 5000
5050 5060 5070 5080 5090 5100 5110
AGGCCAGAGCAAGAGGAGCTGCAGAGAGCCAAAGAGGAGAGCTTGCGCAAGATCTTAAGATGTTAGTTCAT
AGGCCAGAGCAAGAGGAGCTGCAGAGAGCCAAAGAGGAGAGCTTGCGCAAGATCTTAAGATGTTAGTTCAT
5010 5020 5030 5040 5050 5060 5070 5080
5120 5130 5140 5150 5160 5170 5180
TGTACATCTGTAAAGACCAATAGTAAACAAGTCAACAGAACAGATCTGCTGTGGAGACTTGAAGACCA
TGTACATCTGTAAAGACCAATAGTAAACAAGTCAACAGAACAGATCTGCTGTGGAGACTTGAAGACCA
5090 5100 5110 5120 5130 5140 5150
5190 5200 5210 5220 5230 X 5240 5250
GATTTCAGAGCCATATGATTCAGAGAGAGGAGCGCTCAATGATTTTCCACCGATGAGAGACCCGACAT
GATTTCAGAGCCATATGATTCAGAGAGAGGAGCGCTCAATGATTTTCCACCGATGAGAGACCCGACAT
5160 5170 5180 5190 5200
5260 5270 5280
TCTGAGGCTTCAATCGGAGACTGAGAC

US-09-902-432-3 (1-6160)
US-08-978-277A-1 Sequence 1, Application US/08978277A

Sequence 1, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brunbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza


```
1170 1180 1190 1200 1210 1220 1230
2270 2280 2290 2300 2310 2320
CTCCACTTGGAGCATTTAAAGATTAGTCACTCCAAAGAAAAAATCCAAAGTCAAACTGGAGAGAAAGA
CTCCACTTGGAGCATTTAAAGATTAGTCACTCCAAAGAAAAAATCCAAAGTCAAACTGGAGAGAAAGA
1240 1250 1260 1270 1280 1290 1300
2330 2340 2350 2360 2370 2380 2390
AGCC-GAAG-ACCTAG-TGTA-GAGCA-GTTGTCCACTGATCGAACCAGTGAAGAAATCTTGGGTT
AGCC-GAAG-ACCTAG-TGTA-GAGCA-GTTGTCCACTGATCGAACCAGTGAAGAAATCTTGGGTT
2400 2410 2420 2430 2440 2450 2460
TCCATTAAAGAAATCATCCCCGCGACGGCGAAGAAAAAGGCGAGGAA-GCAAGAAACAAGCCACTGTGGA
TCCATTAAAGAAATCATCCCCGCGACGGCGAAGAAAAAGGCGAGGAA-GCAAGAAACAAGCCACTGTGGA
1310 1320 1330 1340 1350 1360 1370 1380
2470 2480 2490 2500 2510 2520 2530
AGACTCAGGGCCAGTGGAGATTAATGAGAGAGACCCCTATATGCCAGCCGTCTGTCTGTGATATTA
AGACTCAGGGCCAGTGGAGATTAATGAGAGAGACCCCTATATGCCAGCCGTCTGTCTGTGATATTA
1460 1470 1480 1490 1500 1510 1520
2540 2550 2560 2570 2580 2590 2600 2610
TGCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
TGCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1530 1540 1550 1560 1570 1580 1590
2620 2630 2640 2650 2660 2670 2680 2690
GTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1600 1610 1620 1630 1640 1650 1660
2690 2700 2710 2720 2730 2740 2750
CAGTGTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
CAGTGTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
1670 1680 1690 1700 1710 1720 1730
2760 2770 2780 2790 2800 2810 2820
AGCCATGCGACCTGTTGAAGAGGTCACTGAAAAGACATCATTTGCAAGAAACTCTGTGCTCAACCCAGAC
AGCCATGCGACCTGTTGAAGAGGTCACTGAAAAGACATCATTTGCAAGAAACTCTGTGCTCAACCCAGAC
1740 1750 1760 1770 1780 1790 1800
2830 2840 2850 2860 2870 2880 2890
GTTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
GTTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
1810 1820 1830 1840 1850 1860 1870 1880
2900 2910 2920 2930 2940 2950 2960 2970
AGCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
AGCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1890 1900 1910 1920 1930 1940 1950
3050 3060 3070 3080 3090 3100 3110
AGAGAGTGTGTGCTGATGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
AGAGAGTGTGTGCTGATGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
2030 2040 2050 2060 2070 2080 2090
3120 3130 3140 3150 3160 3170 3180
GTTGAAAGAGAGTCCAGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG
GTTGAAAGAGAGTCCAGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG
2100 2110 2120 2130 2140 2150 2160
3190 3200 3210 3220 3230 3240 3250
TGAAGAGTGAAGAGAGAGTCCAGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG
TGAAGAGTGAAGAGAGAGTCCAGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG
2170 2180 2190 2200 2210 2220 2230 2240
3260 3270 3280 3290 3300 3310 3320 3330
GCAGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG
GCAGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG
2250 2260 2270 2280 2290 2300 2310
3340 3350 3360 3370 3380 3390 3400
TCCCTGAAGTCAACCGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG
TCCCTGAAGTCAACCGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG
2320 2330 2340 2350 2360 2370 2380
3410 3420 3430 3440 3450 3460 3470
GGCGTGGCCCTTGAAGTCAACCGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA
GGCGTGGCCCTTGAAGTCAACCGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA
2390 2400 2410 2420 2430 2440 2450
3480 3490 3500 3510 3520 3530 3540
AGACATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
AGACATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
2460 2470 2480 2490 2500 2510 2520
3550 3560 3570 3580 3590 3600 3610
GCACTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCACTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
2530 2540 2550 2560 2570 2580 2590 2600
3620 3630 3640 3650 3660 3670 3680 3690
TCCAGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
TCCAGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
2610 2620 2630 2640 2650 2660 2670
3700 3710 3720 3730 3740 3750 3760
TGCAGCCCTGCCCCCTTCTGGC-AAAGACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
TGCAGCCCTGCCCCCTTCTGGC-AAAGACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
2680 2690 2700 2710 2720 2730 2740
3770 3780 3790 3800 3810 3820 3830
AAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
AAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
2750 2760 2770 2780 2790 2800 2810
3840 3850 3860 3870 3880 3890 3900
AAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA
AAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA
2820 2830 2840 2850 2860 2870 2880
3910 3920 3930 3940 3950 3960 3970
AGGAGAGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA
AGGAGAGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA
2890 2900 2910 2920 2930 2940 2950 2960
```

```

3980 3990 4000 4010 4020 4030 4040 4050
AGCAGGAAACAGCGCTCTGAGCATCAGGACCTACGGAGAGCCAGTCTGACATTTGACATGCCAGCT
|||||
AGCAGGAAACAGCGCTCTGAGCATCAGGAAAGCACTACGGGAAAGCAGTCTGACATTTGACATGCCAGCT
2970 2980 2990 3000 3010 3020 3030
4060 4070 4080 4090 4100 4110 4120
CAGAGAGGGGAAGGCACTGGGAAGCTTTGGAGAGCCCTTCTCTCCAGCAGCAAGACAGCAGTTGCA
CAGAGAGGGGAAGGCACTGGGAAGCTTTGGAGAGCCCTTCTCTCCAGCAGCAAGACAGCAGTTGCA
3040 3050 3060 3070 3080 3090 3100
4130 4140 4150 4160 4170 4180 4190
TAGAGGTTCAAGTTCAAGCTTGCAACACAGTCACTCAACACAGCAGAGCTTGGAAGAGTCTATAGAA
TAGAGGTTCAAGTTCAAGCTTGCAACACAGTCACTCAACACAGCAGAGCTTGGAAGAGTCTATAGAA
3110 3120 3130 3140 3150 3160 3170
4200 4210 4220 4230 4240 4250 4260
CGGTTGCTGATTTTCAGAGACAGGTGAAGTCCAGAGTGTAGTGCACACTTATTACAGCTGAGAGTCCT
CGGTTGCTGATTTTCAGAGACAGGTGAAGTCCAGAGTGTAGTGCACACTTATTACAGCTGAGAGTCCT
3180 3190 3200 3210 3220 3230 3240
4270 4280 4290 4300 4310 4320 4330
CTGCAACGGGTGSCCACTGCACTCTTCAGATGCAGAGCAGCGTACCCCTGGGGCCTGAGTCTCAGGCAG
CTGCAACGGGTGSCCACTGCACTCTTCAGATGCAGAGCAGCGTACCCCTGGGGCCTGAGTCTCAGGCAG
3250 3260 3270 3280 3290 3300 3310
4340 4350 4360 4370 4380 4390 4400
AATCCATCCCATCATAGTAACTCTGCTCTGCTGGAAGACCCCTACATCTGACCTTACAGGAGAAATAAGCG
AATCCATCCCATCATAGTAACTCTGCTCTGCTGGAAGACCCCTACATCTGACCTTACAGGAGAAATAAGCG
3330 3340 3350 3360 3370 3380 3390
4420 4430 4440 4450 4460 4470 4480
CATCCAGGAGAGCGATCAGAGGAGAGGACAGCCAGATCTGCTGCTGATGCTGACGGCAAGAGAGTA
CATCCAGGAGAGCGATCAGAGGAGAGGACAGCCAGATCTGCTGCTGATGCTGACGGCAAGAGAGTA
3400 3410 3420 3430 3440 3450 3460
4490 4500 4510 4520 4530 4540 4550
CAGCAATCGAAAAGTCTCAAGGCTGACCTGAGATCTCTGGAACCTTGAGATTAAGAGCAACAAGATTGTC
CAGCAATCGAAAAGTCTCAAGGCTGACCTGAGATCTCTGGAACCTTGAGATTAAGAGCAACAAGATTGTC
3470 3480 3490 3500 3510 3520 3530
4560 4570 4580 4590 4600 4610 4620
TGAACGTCATTTCAGACAGCGCTGACCACTTCGACAGTACAGAAACAGCCCGGAAACTCATGCTTATGATT
TGAACGTCATTTCAGACAGCGCTGACCACTTCGACAGTACAGAAACAGCCCGGAAACTCATGCTTATGATT
3540 3550 3560 3570 3580 3590 3600
4630 4640 4650 4660 4670 4680 4690
CACAGACCCAGGTTCTGCG-ATGCAAGCTT-GACAGCAGGAGCCCAACAGATCTTGCAAAAATGAAGA
CACAGACCCAGGTTCTGCGAATGC--GCTTGACAGCAGGAGCCCAACAGATCTTGCAAAAATGAAGA
3610 3620 3630 3640 3650 3660 3670
4700 4710 4720 4730 4740 4750 4760
TGCCAGATGAACACCCAGTCCGAGCCAGAGAGGACTTGCAAGTCTGACCGTTCTGGAGCGATGGGC
TGCCAGATGAACACCCAGTCCGAGTGAAGCGCGGTGTCAAAGTAAGACTTGGAAGCTGCCT
3680 3690 3700 3710 3720 3730 3740
4770 4780 4790 4800 4810 4820 4830
TCAGCCTGGAATGCTTGCCTGCTGAGTTGAAGCGCGGTGTCAAAGTAAGACTTGGAAGCTGCCT
TCAG-CTCGGAATGCTTGCCTGCTGAGTTGAAGCGCGGTGTCAAAGTAAGACTTGGAAGCTGCCT
3750 3760 3770 3780 3790 3800 3810
4850 4860 4870 4880 4890 4900 4910

```

```

CCTCAACCCAAAGATCCAAAGAGGAGATGCTGCTGATGGCCCTCAGTCCAAAGCTTACCCAGGAGGAGG
|||||
CCTCAACCCAAAGAT-CAAAAGGAGGATGCTGCTGATGGCCCTCAGTCCAAAGCTTACCCAGGAGGAGG-
3830 3840 3850 3860 3870 3880 3890
4920 4930 4940 4950 4960 4970 4980
CAGTGCTCTGGAACCTTAAACCAAGATCCCAAGACCAACCGACCAAGCTAACCCAGGAGGCGAT-C
|||||
CAGTG--TCTGGAACCTTAAACCAAGATCCCAAGACCAACCGACCAAGCTAACCCAGGAGGAGG-
3900 3910 3920 3930 3940 3950
4990 5000 5010 5020 5030 5040 5050
CCCCAAAGTTCAGGTCAGGTCAGGAAAGAAATGCTTACCAAGTCAAGAGAGCAAGCCAGGAGGAGG
CCCCAAAGTTCAGGTCAGGTCAGGAAAGAAATGCTTACCAAGTCAAGAGAGCAAGCCAGGAGGAGG
3960 3970 3980 3990 4000 4010 4020
5060 5070 5080 5090 5100 5110 5120
AGGACTGCGAGGAGCAAGGAGGAGGAGCTGGGAGAAATCCTAAGATGTTAGTCTCATTTGTAATGA
|||||
AGGACTGCGAGGAGCAAGGAGGAGGAGGAGCTGGGAGAAATCCTAAGATGTTAGTCTCATTTGTAATGA
4040 4050 4060 4070 4080 4090 4100
5130 5140 5150 5160 5170 5180 5190
CCAGAAATGTGAAAACCAAGTTCAGAGAAAGATGCTGCTGTTGGGACCTTGAGACCAAGATTTACAGGCCAT
CCAGAAATGTGAAAACCAAGTTCAGAGAAAGATGCTGCTGTTGGGACCTTG-GACCAAGATTTACAGGCCAT
4110 4120 4130 4140 4150 4160 4170
5200 5210 5220 5230 5240 5250 5260
GAGATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GAGATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
4180 4190 4200 4210 4220 4230 4240
5280 5290 5300 5310 5320 5330 5340
GGAGCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT
GGAGCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT
4250 4260 4270 4280 4290 4300 4310
5350 5360 5370 5380 5390 5400
TTC---TAACATTTCTT-CGTT-TCAAGACTGCTTTGATT-TGCCCTTGTATGCC--GTCCGCTATTTCG
TTCCGATTTAAGAGTCTCGTTCCTCAA-CTTGGAAACCAATTTCTGCC---ATACCTAGTTCCACT-TCTCA
4320 4330 4340 4350 4360 4370 4380
5410 5420 5430 5440 5450 5460 5470
GATTTAAG-GTCCCTGCTTCTCAACCTGGAACCAATTCCTGCCATACCTTACCTTCTCAAACTGGAGCA
AACTGGAGCATCTCTTATGATTTT---ATATGATGTTTATGATGATGATGATGATGATGATGATGATG
4390 4400 4410 4420 4430 4440 4450
5480 5490 5500 5510 5520 5530 5540
TCTCTCTTTATGTA---TTTATATGATGTTTATGATGATGATGATGATGATGATGATGATGATGATG
TATT--TTTCTTAAAGTTTAAAGCATGCTTTTGTATGATGATGATGATGATGATGATGATGATGATG
4460 4470 4480 4490 4500 4510 4520
5550 5560 5570 5580 5590 5600
AACGTTT---AAGC-ACATGCTT-TTGTATTATGC-----AATATATAACGGGTGT-GCAGCCATAG-CGA
|||||
ACGTTTGAAGAGTCCAAAGCTCAACTGTAACTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG
4530 4540 4550 4560 4570 4580 4590
5610 5620 5630 5640 5650 5660
CGCT-TTGAAAGCT--CCAAGCTT-CAACTGTAACTGC-AG-----CAAAACAGATAACATTTCTGGCA
GTCCTTTTAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGTGTGTTTCTTCT--A
4600 4610 4620 4630 4640 4650 4660
5670 5680 5690 5700 5710 5720 5730
AGAAAGAGCAAGTCTTTTTTAAAGTTTAACTGATGCTGAGGCTTCTAGTCC---TCTGAAAGTGG

```


US-08-978-277A-2 Sequence 2, Application US/08978277A

Sequence 2, Application US/08978277A

GENERAL INFORMATION:
 APPLICANT: Gelman, Irwin H.
 TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
 STREET: 30 Rockefeller Plaza
 CITY: New York
 STATE: NY
 COUNTRY: USA

ZIP: 10112-0228
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ Version 1.5

CURRENT APPLICATION DATA: US/08/978,277A

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/665,401

FILING DATE: 18-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Richard S

REGISTRATION NUMBER: 26,154

REFERENCE/DOCKET NUMBER: A30558 - 165/34008

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-408-2558

TELEFAX: 212-765-2519

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1346 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

Initial Score = 710 Optimized Score = 1094 Significance = 1.44
 Residue Identity = 89% Matches = 1106 Mismatches = 86
 Gaps = 37 Conservative Substitutions = 0

340 350 360 370 380 390 400
 EQPARDTDOARLSADYKVELPLEDQVGDLEASSEKCAPLATEVDFDKMEAHQEVVAEVHVTVEKTEEE
 MEAHQEVVAEVHVTVEKTEEE
 X 10 20

410 420 430 440 450 460 470 480
 QGGGGAEGGVVVGESLPPEKLAEPQVQPAEELMKSREMCVSGGDHTQTLTDLSPPEKTLPKHPE
 QGGGGAEGGVVVGESLPPEKLAEPQVQPAEELMKSREMCVSGGDHTQTLTDLSPPEKTLPKHPE
 30 40 50 60 70 80 90

490 500 510 520 530 540 550
 GIVSEVEMLSQBRIVQGSPLKLFSSGLKSLGKKQKGGGGDEPGYQHIHTSPESADBEKGES
 GIVSEVEMLSQBRIVQGSPLKLFSSGLKSLGKKQKGGGGDEPGYQHIHTSPESADBEKGES
 100 110 120 130 140 150 160

560 570 580 590 600 610 620
 SASSPEPEPTTCLKGPLEAPQDGAEEGTTSDGKKREGITPWFASFKQWTPKKRVRRPSSDKDEELEK
 SASSPEPEPTTCLKGPLEAPQDGAEEGTTSDGKKREGITPWFASFKQWTPKKRVRRPSSDKDEELEK
 170 180 190 200 210 220 230

630 640 650 660 670 680 690
 VKSATLSSTDSTVSEMQDEVKTGVEQKPEEPKRRVDTSVSWEALICVSSKKRARKASSDDEGPRTLGG
 VKSATLSSTDSTVSEMQDEVKTGVEQKPEEPKRRVDTSVSWEALICVSSKKRARKASSDDEGPRTLGG
 240 250 260 270 280 290 300
 700 710 720 730 740 750 760
 DSHRAEEASKDEAGTDAVPASTQEOQAGSSSPAGSPSEGEVSTWSEFKRLVTPRKSKSKLEK-A
 GQSGRGGQQRQSRRTDAVPASTQEOQAGSSSPAGSPSEGEVSTWSEFKRLVTPRKSKSKLEK-A
 310 320 330 340 350 360 370
 770 780 790 800 810 820
 EDSSV-----EQLSTEIE-----PSREESWVSIKKFIPGRKKRAGDKQEQATVEDSGPVEINEDPNPVA
 GRTLTVGAGCPLRSNRVEKNLGFPLRNS-----PDGGR--KGQMGROEQATVEDSGPVEINEDPNPVA
 380 390 400 410 420 430 440
 830 840 850 860 870 880 890 900
 VPLSEYNAVEREKMEAGNTELPOLLGAVYVSEELSKTLVHTVSVAVIDGTTRAVTSVEERSPSWISASVTEP
 VPLSEYDAVEREKMEAGNTELPOLLGAVYVSEELSKTLVHTVSVAVIDGTTRAVTSVEERSPSWISASVTEP
 450 460 470 480 490 500 510
 910 920 930 940 950 960 970
 LEHTAGEAMPVVEEVTEKDIIAETPVLTLPEGKADHDMVTSEVDFTSSEAVTATETSEALRTEEVTEAS
 LEHTAGEAMPVVEEVTEKDIIAETPVLTLPEGKADHDMVTSEVDFTSSEAVTATETSEALRTEEVTEAS
 520 530 540 550 560 570 580
 980 990 1000 1010 1020 1030 1040
 GAEETDMVSAYSQLTSDPTTEATPVQEVESGVLDTTEERQTOAILQAVADKVKESQVPAQTQVQRTG
 GAEETDMVSAYSQLTSDPTTEATPVQEVESGVLDTTEERQTOAILQAVADKVKESQVPAQTQVQRTG
 590 600 610 620 630 640 650
 1050 1060 1070 1080 1090 1100 1110
 SKALEKVEEVEDSEVLASEKEKDMVMPKGPVQVQEAHQAQSETQATPESLEVEPVDVADHVAATQVQV
 SKALEKVEEVEDSEVLASEKEKDMVMPKGPVQVQEAHQAQSETQATPESLEVEPVDVADHVAATQVQV
 660 670 680 690 700 710 720 730
 1120 1130 1140 1150 1160 1170 1180
 LQQLMEQAVAPESSETLTDSETNGSTPLADSDTADGTQODETIDSDSKATAAVROSQVTEEEAATAQKEEP
 LQQLMEQAVAPESSETLTDSETNGSTPLADSDTADGTQODETIDSDSKATAAVROSQVTEEEAATAQKEEP
 1190 1200 1210 1220 1230 1240 1250 1260
 STLNNVPAQEEHGEPRGRDLVLEPTQOELTAAAPVLAKEVQGEVDWLDGKVKESQVVFHSGPNSQK
 STLNNVPAQEEHGEPRGRDLVLEPTQOELTAAAPVLAKEVQGEVDWLDGKVKESQVVFHSGPNSQK
 810 820 830 840 850 860 870
 1270 1280 1290 1300 1310 1320 1330
 AADVTYDSVMGVAGQEKESTEVOSLSLEEGEMETDVEKRETKPEQVSEEGEQTAAPEHRYGKPV
 AADVTYDSVMGVAGQEKESTEVOSLSLEEGEMETDVEKRETKPEQVSEEGEQTAAPEHRYGKPV
 880 890 900 910 920 930 940
 1340 1350 1360 1370 1380 1390 1400
 TLDMPSSRGKALGSLGGSPSLPDQDKACIEVQSLDVTVTQTAEEAVEKVIETVITSETGESPCVGAHL
 TLDMPSSRGKALGSLGGSPSLPDQDKACIEVQSLDVTVTQTAEEAVEKVIETVITSETGESPCVGAHL
 950 960 970 980 990 1000 1010
 1410 1420 1430 1440 1450 1460 1470
 LPAEKSSATGGHWTLOHARDTVPPLGPESQAESIPITVTPAPESTLHPDLQGISASORSSEEDKPDAGPD
 LPAEKSSATGGHWTLOHARDTVPPLGPESQAESIPITVTPAPESTLHPDLQGISASORSSEEDKPDAGPD
 1020 1030 1040 1050 1060 1070 1080 1090

1480 1490 1500 1510 1520 1530 1540
ADGESIAIEVLAEPEILIELESKSNKIYINVIQTAVDQFARTETAPETAYISQTVAPACRIDSREPNC
ADGESIAIDIKVLAEPEILIELESKSNKIYINVIQTAVDQFARTETAPETAYISQTVAPACRIDSREPNC
ADGESIAIDIKVLAEPEILIELESKSNKIYINVIQTAVDQFARTETAPETAYISQTVAPACRIDSREPNC
1100 1110 1120 1130 1140 1150 1160
WTGKAKAKMKHPVQPREDLQVLTVLEKMAQPRKCLPRLQLKA-----PVSK X
1170 1180 1190 1200 1210 1220
WTGKAKAKMKHPVQPREDLQVLTVLEKMAQPRKCLPRLQLKA-----PVSK X
POLQSLAQAENAVSGNLTRESPTNGPKLTEREC
1230 1240 1250 1260

3. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 18, Application US/08978277A

Sequence 18, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 23 Optimized Score = 23 Significance = -0.28
Residue Identity = 100% Matches = 23 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

700 710 720 730 740 750 760
PRTLGDSHRAEASKDEAGTDAVPASTQEDQAGSSSPBAGSPBEGSGVSTWESFKRLVTPRKKSXSK
|||||
EGVSTWESFKRLVTPRKKSXSK
X 10 20

X 770 780 790 800 810
LBEAKEDSVYQLSTELIEPBRRESWYSIKKFIPIGRKKRADGKQEQATVVD
L
X

4. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 17, Application US/08978277A

Sequence 17, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 23 Optimized Score = 23 Significance = -0.28
Residue Identity = 100% Matches = 23 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

550 560 570 580 590 X 600 610
PESADEQKGESSASSPPEPPEPTTCLKGPLEAPQDGEABEGTTSDEGKKGITPPWASFKKMTTPKKRVRRP
|||||
EGITPPWASFKKMTTPKKRVRRP
X 10 20

X 620 630 640 650 660
SESDKEBELKVKSATLSTSTVSEMDVKTGGEQKPEEPKRRVDTSV
|
S
X

5. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 16, Application US/08978277A

Sequence 16, Application US/08978277A

GENERAL INFORMATION: 05/06597827/A

APPLICANT: Geiman, Irwin H.

TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: NY

COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978.277A

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S.
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 -
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 16:

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

```
Initial Score      = 23  Optimized Score = 23  Significance = -0.28
Residue Identity = 100%  Matches      = 23  Mismatches   = 0
Gaps              = 0    Conservative Substitutions
```

EMCVSGDHTQLTDLSPEKKTLPH**PCEI**VSEVNLSSQERIKVG**SFLKLFSSSGLKL**GSKKGKGGG
KLFSSSGLKLGSKKGKGGG
X X 10 20

X 530 540 550 560 570
GGDEEPGEYQHTSPESADEKQGSASSPEETTCLEKGPLEAPQD
— G X

6. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 15, Application US/08978277A

Sequence 15, Application US/08978277A

SEQUENCE 157, APPLICATION NO. 06/978277/A
GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza

CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978, 277A

FILING DATE: 514
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/665,401
 FILING DATE: 18-JUN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Richard S
 REGISTRATION NUMBER: 26,154
 REFERENCE/DOCKET NUMBER: A30558 - 165/34008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-408-2558
 TELEFAX: 212-765-2519

INFORMATION FOR SEQ ID NO: 15:

INFORMATION FOR SEQU ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRAINEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

```
Initial Score      = 13  Optimized Score = 22  Significance = -0.31
Residue Identity = 76%  Matches      = 23  Mismatches  = 0
Gaps              = 7    Conservative Substitutions
```

230 240 250 260 270 280 290 300
TLAQEQSSTEIPLQAEQSDQAABEAKDEGEKEFTKSPESPVSINETSSFKKFTTHGWAGWRKITS
ETSSFKKFTTHG-----TS
X 10

310 320 330 340 350
FKKSKEDDLETAEKKEQAEKVDEEKEKTEPASEQEPAEDTDOARLSADYEKVEL
|||||
FKKSKEDD
20 X

7. US-09-902-432-4 (1-1596)
US-08-978-277A-7 Sequence 7, Application US/08978277A

Sequence 7, Application US/09878277A
GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5

Initial Score = 7 Optimized Score = 7 Significance = -0.32
Residue Identity = 100% Matches = 7 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

250 260 270 280 290 X 300 310
QAESDQAABEAKDEGEKQKPTKSPSPVNSSETTSFKFFTHGAGWRKKTSTFKSKDEDDLETAE
|||||
WAGWRKK
X X

320 330 340
KRKEQAEKVDSEKEKTEPASBEQEPADTDQAR

10. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 13, Application US/08978277A

Sequence 13, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 7 Optimized Score = 8 Significance = -0.32
Residue Identity = 47% Matches = 8 Mismatches = 9
Gaps = 0 Conservative Substitutions = 0

460 470 480 490 500 X 510 520
EMCVSGDHTQLTDLSPBEKTLPKHPGIVSEVMSLSQRIKVGSPPLKLFSSSGLKKLSGKKQKRG
|||||
KKFSSKKPKLSGFSR
X 10 X

530 540 550 560 570
GGDEEPGEYQHHTSPESADEQKGSASSPPEETTCLEKGP

11. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 12, Application US/08978277A

Sequence 12, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 7 Optimized Score = 10 Significance = -0.32
Residue Identity = 47% Matches = 11 Mismatches = 9
Gaps = 3 Conservative Substitutions = 0

460 470 480 490 500 X 510 520
EMCVSGDHTQLTDLSPBEKTLPKHPGIVSEVMSLSQRIKVGSPPLKLFSSSGLKKLSG---KKQKGG
|||||
KKFSSKKPKLSGFSFKGKKE
X 10 20

X 530 540 550 560 570
RGGGDEEPGEYQHHTSPESADEQKGSASSPPEETTCLEKGP
A
X

12. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 11, Application US/08978277A

Sequence 11, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

COMPUTER READABLE FORM:

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 5 Optimized Score = 5 Significance = -0.33
Residue Identity = 50% Matches = 5 Mismatches = 5
Gaps = 0 Conservative Substitutions = 0

250 260 270 280 290 300 320
ABEEAKGEKEKTEPASEQRPEDTDQARLSADYKVEL
|||||
MGCIKSKEDK
X 10

330 340 350
EKVDEBEKEKTEPASEQRPEDTDQARLSADYKVEL

15. US-09-902-432-4 (1-1596)
US-08-978-277A-5 Sequence 5, Application US/08978277A

Sequence 5, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 4 Optimized Score = 4 Significance = -0.33
Residue Identity = 40% Matches = 4 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

710 720 730 740 750 760 770
ASKDKRAGTDAVPASTQEQDQAGSSSPGAGSPGEGGVSTWESFKRLVTRKKSLSKLEAKDSSVEQL
|||||
MGSSKSKPKD
X 10

780 790 800 810
STEIEPSREESNVSIKKFIPGRKKRADGKQEQATVED

16. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 14, Application US/08978277A

Sequence 14, Application US/08978277A

GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:

Initial Score = 4 Optimized Score = 4 Significance = -0.33
Residue Identity = 20% Matches = 4 Mismatches = 16
Gaps = 0 Conservative Substitutions = 0

690 700 710 720 730 740 750 X

SSSDDECGPTLGGDSHRAEASKDKAGTDAVPASTQEDQAGSSSPSEGEVSTWESFKRLVT

|||
KRRMKAFIAVSAARFKC
X 10 20

760 770 780 790 800
PRKSKSLKEKAEDSSVEQLSTIEFRESRWSWISIKKFIPIGRKKRA

17. US-09-902-432-4 (1-1596)
US-08-978-277A-9 Sequence 9, Application US/08978277A

Sequence 9, Application US/08978277A

GENERAL INFORMATION:

APPLICANT: Gelman, Irwin H.

TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10112-0228

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978,277A

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/665,401

FILING DATE: 18-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Richard S

REGISTRATION NUMBER: 26,154

REFERENCE/DOCKET NUMBER: A30558 - 165/34008

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-408-2558

TELEFAX: 212-765-2519

TELEX:

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

Initial Score = 3 Optimized Score = 4 Significance = -0.33
Residue Identity = 40% Matches = 4 Mismatches = 6
Gaps = 0 Conservative Substitutions = 0

580 590 600 610 620 630 640
DGEAEGTTSDEGKEREGITTPWASFKKVTPKRRVRPSSESDKEBELEKVSATLSSTDSTVSEMODEVKTV
|||
MGCTLSAEDK
X 10

650 660 670 680
GGEOKPEEPKRRVDTSVSWALICVGSSEKRRARAKASS

18. US-09-902-432-4 (1-1596)
US-08-978-277A-1 Sequence 10, Application US/08978277A

Sequence 10, Application US/08978277A

GENERAL INFORMATION:

APPLICANT: Gelman, Irwin H.

TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10112-0228

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978,277A

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/665,401

FILING DATE: 18-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Richard S

REGISTRATION NUMBER: 26,154

REFERENCE/DOCKET NUMBER: A30558 - 165/34008

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-408-2558

TELEFAX: 212-765-2519

TELEX:

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

Initial Score = 2 Optimized Score = 2 Significance = -0.34
Residue Identity = 20% Matches = 2 Mismatches = 8
Gaps = 0 Conservative Substitutions = 0

560 570 580 590 600 610 620
SASSPEPEETTCLEKGPLAPDGEAEBGTTSDEGKKRGITTPWASFKKVTPKRRVRPSSESDKEBELEK
|||
MLCCMRITKQ
X 10

630 640 650 660
VKSATLSSTDSTVSEMODEVKTVGEOKPEEPKRRVDT

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:13:30 ; Search time 11503 Seconds
(without alignments)
13464.116 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cgggcccggggagtagag.....aataaacatatttgactac 6160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main.*
1: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081 COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082 COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083 COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084 COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085 COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087 COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088 COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089 COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090 COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091 COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092 COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093 COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US094 COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US095A COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US095B COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095C COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095D COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US096A COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US096B COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096C COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096D COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096E COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US097A COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US097B COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097C COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US098A COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US098B COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098C COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US099A COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US099B COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US099C COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US099D COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US100A COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6023 COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6026 COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6160	100.0	6160	1	PCT-US97-06830-3	Sequence 3, Appli
2	6160	100.0	6160	10	US-08-665-401-3	Sequence 3, Appli
3	6160	100.0	6160	34	US-09-902-432-3	Sequence 3, Appli
4	5204.2	84.5	5236	1	PCT-US02-16173A-2063	Sequence 2063, Ap
5	5204.2	84.5	5236	41	US-10-191-803-279	Sequence 279, App
6	5200	84.4	5200	13	US-08-978-277A-3	Sequence 3, Appli
7	4831.4	78.4	5134	1	PCT-US97-06830-1	Sequence 1, Appli
8	4831.4	78.4	5134	10	US-08-665-401-1	Sequence 1, Appli
9	4831.4	78.4	5134	34	US-09-902-432-1	Sequence 1, Appli
10	4701.4	76.3	5074	13	US-08-978-277A-1	Sequence 1, Appli
11	4548.6	73.8	6199	80	US-06-360-207-12210	Sequence 12210, A
12	2181.2	35.4	7651	30	US-09-770-174-4018	Sequence 4018, Ap
13	2178	35.4	6803	71	US-60-278-232-10143	Sequence 10143, A
14	2178	35.4	6607	68	US-60-245-228-517	Sequence 517, App
15	2172.8	35.3	6632	76	US-60-324-185-30580	Sequence 30580, A
16	2160.4	35.1	8660	25	US-09-644-867-6364	Sequence 6364, Ap
17	2160.4	35.1	8660	25	US-09-644-869-9117	Sequence 9117, Ap
18	2160.4	35.1	8660	25	US-09-652-911-10117	Sequence 10117, A
19	2160.4	35.1	8660	25	US-09-652-913-10715	Sequence 10715, A
20	2160.4	35.1	8660	25	US-09-652-918-8758	Sequence 8758, Ap
21	2160.4	35.1	8660	26	US-09-667-760-2000	Sequence 2000, Ap

22 2160.4 35.1 8660 27 US-09-699-997-11667
 23 2160.4 35.1 8660 27 US-09-699-999-6654
 24 2160.4 35.1 8660 28 US-09-700-000-5679
 25 2160.4 35.1 8660 28 US-09-710-280-2315
 26 2160.4 35.1 8660 28 US-09-710-286-3428
 27 2160.4 35.1 8660 28 US-09-716-472-4599
 28 2160.4 35.1 8660 28 US-09-716-953-2245
 29 2160.4 35.1 8660 29 US-09-726-789-3701
 30 2160.4 35.1 8660 29 US-09-726-802-2407
 31 2160.4 35.1 8660 29 US-09-726-805-1797
 32 2160.4 35.1 8660 29 US-09-726-806-5392
 33 2159.6 35.1 8492 23 US-09-606-776-5120
 34 2144.4 34.8 6608 16 US-09-220-132-58
 35 2144.4 34.8 6608 29 US-09-738-877-1
 36 2144.4 34.8 6608 33 US-09-880-107-3439
 37 2144.4 34.8 6608 41 US-10-155-653-58
 38 2108.8 34.2 8869 71 US-60-324-185-30584
 39 2106.2 34.2 8869 71 US-60-278-232-10145
 40 2106.2 34.2 8888 65 US-60-213-359-7019
 41 2094.6 34.0 8877 61 US-60-172-360-24882
 42 2055.6 33.4 8243 17 US-09-397-022-4579
 43 2055.6 33.4 8243 18 US-09-404-549-4757
 44 2055.6 33.4 8243 18 US-09-404-549A-4757
 45 2055.6 33.4 8243 18 US-09-428-151A-10174

Sequence 11667, A
 Sequence 6654, Ap
 Sequence 5679, Ap
 Sequence 2315, Ap
 Sequence 3428, Ap
 Sequence 4599, Ap
 Sequence 2245, Ap
 Sequence 3701, Ap
 Sequence 2407, Ap
 Sequence 1797, Ap
 Sequence 5392, Ap
 Sequence 5120, Ap
 Sequence 58, Ap
 Sequence 1, Appl
 Sequence 3439, Ap
 Sequence 58, Appl
 Sequence 10145, A
 Sequence 7019, Ap
 Sequence 2488, A
 Sequence 4579, Ap
 Sequence 4757, Ap
 Sequence 10174, A

ALIGNMENTS

RESULT 1
 PCT-US97-06830-3
 Sequence 3, Application PC/TUS9706830
 GENERAL INFORMATION:
 APPLICANT: Gelman, Irwin H.
 TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
 STREET: 30 Rockefeller Plaza
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10112-0228
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastrSEO Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US97/06830
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/665,401
 FILING DATE: 18-JUN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Richard S
 REGISTRATION NUMBER: 26,154
 REFERENCE/DOCKET NUMBER: A30558 - 165/34008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-408-2558
 TELEFAX: 212-765-2519
 TELEX:
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6160 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:

ORIGINAL SOURCE:
 PCT-US97-06830-3

Query Match 100.0%; Score 6160; DB 1; Length 6160;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 6160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCCGGCGGAGTAAAGCCATGACCTGAGCGGCGGACGATTCACCGAGAGC 60
 1 CGGGCCGGCGGAGTAAAGCCATGACCTGAGCGGCGGAGTTCACCGAGAGC 60
 QY 61 GGAGCCCGGAGCAGCCCGGCGGAGCAGCAGCCGAGCAGCTGCTCAGTGGCCATG 120
 1 GGAGCCCGGAGCAGCCCGGCGGAGCAGCAGCCGAGCAGCTGCTCAGTGGCCATG 120
 Db 61 GGAGCCCGGAGCAGCCCGGCGGAGCAGCAGCCGAGCAGCTGCTCAGTGGCCATG 120
 QY 121 GGCCCGGAGCTGAAGCCTCGGAGCAGCTGAGAGCCCGCGAGCGGACCCCGCACA 180
 121 GGCCCGGAGCTGAAGCCTCGGAGCAGCTGAGAGCCCGCGAGCGGACCCCGCACA 180
 Db 121 GGCCCGGAGCTGAAGCCTCGGAGCAGCTGAGAGCCCGCGAGCGGACCCCGCACA 180
 QY 181 AGCTCCCAAGAGATGCGCAGCTGCTTCTGTCACAGCGCTAGCTGAACAGAGATG 240
 181 AGCTCCCAAGAGATGCGCAGCTGCTTCTGTCACAGCGCTAGCTGAACAGAGATG 240
 Db 181 AGCTCCCAAGAGATGCGCAGCTGCTTCTGTCACAGCGCTAGCTGAACAGAGATG 240
 QY 241 TCCATGTCGAAGAGAAACGAGAGGGGAGAGAGAAAGTGTGATGAGATGTTG 300
 241 TCCATGTCGAAGAGAAACGAGAGGGGAGAGAGAAAGTGTGATGAGATGTTG 300
 Db 241 TCCATGTCGAAGAGAAACGAGAGGGGAGAGAGAAAGTGTGATGAGATGTTG 300
 QY 301 GACAGCGAGAGTCAGAAATGTGAGAGAAAGACCGAGTTGAAGAAATGGCGCACT 360
 301 GACAGCGAGAGTCAGAAATGTGAGAGAAAGACCGAGTTGAAGAAATGGCGCACT 360
 Db 301 GACAGCGAGAGTCAGAAATGTGAGAGAAAGACCGAGTTGAAGAAATGGCGCACT 360
 QY 361 CCAAGCTGTTGAAGATATCAAAAGATGGCGAGAGAGACATCAAGAAATTAATGAC 420
 361 CCAAGCTGTTGAAGATATCAAAAGATGGCGAGAGAGACATCAAGAAATTAATGAC 420
 Db 361 CCAAGCTGTTGAAGATATCAAAAGATGGCGAGAGAGACATCAAGAAATTAATGAC 420
 QY 421 AGATCCCTGCTTCAGAAACATGTGGAAGAAATGTGACAGCTGAGTCCAGGCTA 480
 421 AGATCCCTGCTTCAGAAACATGTGGAAGAAATGTGACAGCTGAGTCCAGGCTA 480
 Db 421 AGATCCCTGCTTCAGAAACATGTGGAAGAAATGTGACAGCTGAGTCCAGGCTA 480
 QY 481 ATGATGTTGGCTTCAGAAAGATATTAATTTGTTGTTTAAATTCACGGTGAAGAG 540
 481 ATGATGTTGGCTTCAGAAAGATATTAATTTGTTGTTTAAATTCACGGTGAAGAG 540
 Db 481 ATGATGTTGGCTTCAGAAAGATATTAATTTGTTGTTTAAATTCACGGTGAAGAG 540
 QY 541 ATAAATATGAAGATATGATATCTGTCATCTCATCTGTCAGAAAGATGAAGCGAAG 600
 541 ATAAATATGAAGATATGATATCTGTCATCTCATCTGTCAGAAAGATGAAGCGAAG 600
 Db 541 ATAAATATGAAGATATGATATCTGTCATCTCATCTGTCAGAAAGATGAAGCGAAG 600
 QY 601 GGGCAGAGCCTCTGTGAGAGCTGAGAGCCAGAGAGCCAGTGTGAGACTCCGTCG 660
 601 GGGCAGAGCCTCTGTGAGAGCTGAGAGCCAGAGAGCCAGTGTGAGACTCCGTCG 660
 Db 601 GGGCAGAGCCTCTGTGAGAGCTGAGAGCCAGAGAGCCAGTGTGAGACTCCGTCG 660
 QY 661 GAGAGTCAGATCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAAAGCAAGGCCACC 720
 661 GAGAGTCAGATCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAAAGCAAGGCCACC 720
 Db 661 GAGAGTCAGATCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAAAGCAAGGCCACC 720
 QY 721 TGAAGCAAGAAAGAGAGCAAGCAAGAAATCCCTTCAAGCCGAATCTGATCAAGCGCTG 780
 721 TGAAGCAAGAAAGAGAGCAAGCAAGAAATCCCTTCAAGCCGAATCTGATCAAGCGCTG 780
 Db 721 TGAAGCAAGAAAGAGAGCAAGCAAGAAATCCCTTCAAGCCGAATCTGATCAAGCGCTG 780
 QY 781 AGAAGAAAGCCAAAGATGAAGAGAGAGAAAGCAAGAAAGAGCCAGCAAGTCCCGAG 840
 781 AGAAGAAAGCCAAAGATGAAGAGAGAGAAAGCAAGAAAGAGCCAGCAAGTCCCGAG 840
 Db 781 AGAAGAAAGCCAAAGATGAAGAGAGAGAAAGCAAGAAAGAGCCAGCAAGTCCCGAG 840
 QY 841 AATCCCGAGAGCCAGTCAAGTGAAGCAATCTTCTTCAAGAGTCTTCACTC 900
 841 AATCCCGAGAGCCAGTCAAGTGAAGCAATCTTCTTCAAGAGTCTTCACTC 900
 Db 841 AATCCCGAGAGCCAGTCAAGTGAAGCAATCTTCTTCAAGAGTCTTCACTC 900
 QY 901 ACGTTGGGCGGCTGGGCAAGAAACAGCTTCAAGAAATCAAAAAGAGATATCGG 960
 901 ACGTTGGGCGGCTGGGCAAGAAACAGCTTCAAGAAATCAAAAAGAGATATCGG 960
 Db 901 ACGTTGGGCGGCTGGGCAAGAAACAGCTTCAAGAAATCAAAAAGAGATATCGG 960
 QY 961 AAATGCGGAG 1020
 961 AAATGCGGAG 1020
 Db 961 AAATGCGGAG 1020

Db 961 AAACCTGCCAGAGAGAAAGGAGCAGAGGCGAGAAAAGTAGACGAGGAAGAAAGGAA 1020
Qy 1021 AGACAGAGCCAGCCTCGGAGAGCAGGAGCCGCGAGAGACACAGACAGAGCCAGGTTGT 1080
Db 1021 AGACAGAGCCAGCCTCGGAGAGCAGGAGCCGCGAGAGACACAGACAGAGCCAGGTTGT 1080
Qy 1081 CAGCAGACTACGAGAGGTGAGCTGCTTTGGCAACGGAAGTGTGATGAGAAGATGGAAGCCC 1140
Db 1081 CAGCAGACTACGAGAGGTGAGCTGCTTTGGCAACGGAAGTGTGATGAGAAGATGGAAGCCC 1140
Qy 1141 CGTCAGAGGAGAGTGTGCTCTTTGGCAACGGAAGTGTGATGAGAAGATGGAAGCCC 1200
Db 1141 CGTCAGAGGAGAGTGTGCTCTTTGGCAACGGAAGTGTGATGAGAAGATGGAAGCCC 1200
Qy 1201 ACCAAGAGTTGTTGACAGAGTCCACGTGACACCGGTGAGAGAACAGAGAGGAGCAGG 1260
Db 1201 ACCAAGAGTTGTTGACAGAGTCCACGTGACACCGGTGAGAGAACAGAGAGGAGCAGG 1260
Qy 1261 GAGGAGGAGAGAGGCTGAAGGGGCGGTGTTGAGAGAACAGAGAGGAGGAGCAGG 1320
Db 1261 GAGGAGGAGAGAGGCTGAAGGGGCGGTGTTGAGAGAACAGAGAGGAGGAGCAGG 1320
Qy 1321 CTGAGAAACTCGCTGAGCCCGCAGGAGTCCCGCAGGAGCTGAGCCTGCTGAGGAGCTGA 1380
Db 1321 CTGAGAAACTCGCTGAGCCCGCAGGAGTCCCGCAGGAGCTGAGCCTGCTGAGGAGCTGA 1380
Qy 1381 TGAAGAGCAGAGAGATGTGCTCTGAGAGGAGACCACTCAACTGACAGACCTAAGTC 1440
Db 1381 TGAAGAGCAGAGAGATGTGCTCTGAGAGGAGACCACTCAACTGACAGACCTAAGTC 1440
Qy 1441 CTGAAGAGAGACCTGCCCAACACCCAGAGAGGATGTCAGTGAGGTGAGATGCTGT 1500
Db 1441 CTGAAGAGAGACCTGCCCAACACCCAGAGAGGATGTCAGTGAGGTGAGATGCTGT 1500
Qy 1501 CCTCTCAGGAAGAACTCAAGGTACAGGAGTCCCTTGAAGAACTCTTCAAGTACTGAC 1560
Db 1501 CCTCTCAGGAAGAACTCAAGGTACAGGAGTCCCTTGAAGAACTCTTCAAGTACTGAC 1560
Qy 1561 GCTTAAAGAACTGTCTGGGAAGAGCAGAGAGGGGAAACGAGGAGGTGGGGAGACGAAG 1620
Db 1561 GCTTAAAGAACTGTCTGGGAAGAGCAGAGAGGGGAAACGAGGAGGTGGGGAGACGAAG 1620
Qy 1621 AGCCTGGAGAACTACCAACATTCACACCGAATCCCGAGAGAGTGTGATGAGCAGAAG 1680
Db 1621 AGCCTGGAGAACTACCAACATTCACACCGAATCCCGAGAGAGTGTGATGAGCAGAAG 1680
Qy 1681 GAGAGAGCTCTGCTGCTGCTCCCGAGGAGCCTGAGGAGACCACTGTCTGAGAAAGGGC 1740
Db 1681 GAGAGAGCTCTGCTGCTGCTCCCGAGGAGCCTGAGGAGACCACTGTCTGAGAAAGGGC 1740
Qy 1741 CGCTGGAAGCACCCAGGATGGGAAGCTGAGGAAGGAACCTACTTCCGATGGAGAGAAGA 1800
Db 1741 CGCTGGAAGCACCCAGGATGGGAAGCTGAGGAAGGAACCTACTTCCGATGGAGAGAAGA 1800
Qy 1801 AGAGAGAGGAGATCACTCCCTGGGATCTCTTCAAAAAGATGGTGACACCCCAAGAAACGG 1860
Db 1801 AGAGAGAGGAGATCACTCCCTGGGATCTCTTCAAAAAGATGGTGACACCCCAAGAAACGG 1860
Qy 1861 TCCGAAGACCTTCTGAGAGTGCACAGGAGGAGAGCTGGAAGGTCAAGAGCCCACT 1920
Db 1861 TCCGAAGACCTTCTGAGAGTGCACAGGAGGAGAGCTGGAAGGTCAAGAGCCCACT 1920
Qy 1921 TGTCTCCACTGATAGCACAGTGTGAGAAATGCAAGATGAAAGTCAAACTGTGTGAGG 1980
Db 1921 TGTCTCCACTGATAGCACAGTGTGAGAAATGCAAGATGAAAGTCAAACTGTGTGAGG 1980
Qy 1981 AACAAAAGCCAGAGAACCAAGCGTAGGTGGATACCTTCAAGTGTCTTGGGAAGCACTGA 2040
Db 1981 AACAAAAGCCAGAGAACCAAGCGTAGGTGGATACCTTCAAGTGTCTTGGGAAGCACTGA 2040
Qy 2041 TTTGTGTCGGATCATCCNAGAGAGAGCAAGAGGATCTCTTCAAGTGTGATGAGGAG 2100
Db 2041 TTTGTGTCGGATCATCCNAGAGAGAGCAAGAGGATCTCTTCAAGTGTGATGAGGAG 2100

Qy 2101 GGCCAAAGGACACTGGGAGGGGACAGTCAAGAGCAGAGGAGGCCAGCAAAAGACAAAG 2160
Db 2101 GGCCAAAGGACACTGGGAGGGGACAGTCAAGAGCAGAGGAGGCCAGCAAAAGACAAAG 2160
Qy 2161 CCGGAAACAGAGCGTGTCTCCAGACACCCAGGAGCAGGACCAAGCGCAAGGAGTTTCT 2220
Db 2161 CCGGAAACAGAGCGTGTCTCCAGACACCCAGGAGCAGGACCAAGCGCAAGGAGTTTCT 2220
Qy 2221 CACCCGAGCCAGGGGAGGCCCTTCCGAAGGGAGGTGTCTCCACTTGGGAGTCATTTA 2280
Db 2221 CACCCGAGCCAGGGGAGGCCCTTCCGAAGGGAGGTGTCTCCACTTGGGAGTCATTTA 2280
Qy 2281 AAAGATTAGTCACTCCAAAGAAAAAATCCAAAGTCAAATCTGGAGAGAAAAAGCACT 2340
Db 2281 AAAGATTAGTCACTCCAAAGAAAAAATCCAAAGTCAAATCTGGAGAGAAAAAGCACT 2340
Qy 2341 CTAGTGTAGAGCAGTGTCTCACTGAGATCGAACCGAGTAGAGAAATCTTTGGTTTCCA 2400
Db 2341 CTAGTGTAGAGCAGTGTCTCACTGAGATCGAACCGAGTAGAGAAATCTTTGGTTTCCA 2400
Qy 2401 TTAAGAAATTCATCCCGGACGCGGAAGAAAAAGGGGAGAGCAAGCAAGCAAGCA 2460
Db 2401 TTAAGAAATTCATCCCGGACGCGGAAGAAAAAGGGGAGAGCAAGCAAGCAAGCA 2460
Qy 2461 CTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCCCTTAATGTCTCCAGCCGTG 2520
Db 2461 CTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCCCTTAATGTCTCCAGCCGTG 2520
Qy 2521 TGCTCTGTCTGAGTAAATGTCAGTGGAGAGGAGAGATGGAAGCCAGGGGAAATACGG 2580
Db 2521 TGCTCTGTCTGAGTAAATGTCAGTGGAGAGGAGAGATGGAAGCCAGGGGAAATACGG 2580
Qy 2581 AGCTGCCCCAGCTGTCTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTC 2640
Db 2581 AGCTGCCCCAGCTGTCTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTC 2640
Qy 2641 ACACCTGTGAGTGTGCGAGTCAITGATGGGACCCAGGGCAGTCACCCAGTGTGGAAGAGCGGT 2700
Db 2641 ACACCTGTGAGTGTGCGAGTCAITGATGGGACCCAGGGCAGTCACCCAGTGTGGAAGAGCGGT 2700
Qy 2701 CTCTCTGCTGGATATCCGCTTCGTTACAGAACTCTTTGAAACACACAGCGGGAGAGCCA 2760
Db 2701 CTCTCTGCTGGATATCCGCTTCGTTACAGAACTCTTTGAAACACACAGCGGGAGAGCCA 2760
Qy 2761 TGCCACCTGTGAAGAGGTCACTGAAAAGACATCATTTGAGAGAACTCTCTGTGCTCA 2820
Db 2761 TGCCACCTGTGTGAAGAGGTCACTGAAAAGACATCATTTGAGAGAACTCTCTGTGCTCA 2820
Qy 2821 CCCAGACGTTTACCAGAGGGTAAAGATGCCATGACGACATGGTCAACCCAGTGAAGTGGATT 2880
Db 2821 CCCAGACGTTTACCAGAGGGTAAAGATGCCATGACGACATGGTCAACCCAGTGAAGTGGATT 2880
Qy 2881 TCACCTCAGAACTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTAAGAGATT 2940
Db 2881 TCACCTCAGAACTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTAAGAGATT 2940
Qy 2941 CCGAAGCATCGGGGCGGAGAGACCAAGAGTGTGTCCGAGTTTCCAGCTGACTG 3000
Db 2941 CCGAAGCATCGGGGCGGAGAGACCAAGAGTGTGTGTCCGAGTTTCCAGCTGACTG 3000
Qy 3001 ACTCCCGAGACACACAGAGAGGAGCCACCCAGTTCAGGAGGTAGAGAGTGGTGTGCTAG 3060
Db 3001 ACTCCCGAGACACACAGAGAGGAGCCACCCAGTTCAGGAGGTAGAGAGTGGTGTGCTAG 3060
Qy 3061 ATACAGAAAGAGGAGCGCCAGACGAGGCCATCTCTCAAGCGGTTGCGAGCAAGGTGA 3120
Db 3061 ATACAGAAAGAGGAGCGCCAGACGAGGCCATCTCTCAAGCGGTTGCGAGCAAGGTGA 3120
Qy 3121 AAGAGGAGTCCAGGTCCTGCAACCCAGAGTGTGAGAGAACGGGGTCAAAAGCACTGG 3180
Db 3121 AAGAGGAGTCCAGGTCCTGCAACCCAGAGTGTGAGAGAACGGGGTCAAAAGCACTGG 3180

QY	3181	AGAAAGTTGAGAGAGTGAAGAGAGACTCCGAAAGTGTGGCTTCGGAGAAAGAAAGAGACG	3240
Db	3181	AGAAAGTTGAGAGAGTGAAGAGAGACTCCGAAAGTGTGGCTTCGGAGAAAGAAAGAGACG	3240
QY	3241	TTTAAGCCGAAAGGACCCGTGCAGGAAAGCTGGAGCTGAGACATTTGACACAGGAGCTCTGAGA	3300
Db	3241	TTTAAGCCGAAAGGACCCGTGCAGGAAAGCTGGAGCTGAGACATTTGACACAGGAGCTCTGAGA	3300
QY	3301	CTGACACAGGCTACTCCAGAGAGGCTTGAAGTTCTTGAATCTACGCGCAGATGAGACCATG	3360
Db	3301	CTGACACAGGCTACTCCAGAGAGGCTTGAAGTTCTTGAATCTACGCGCAGATGAGACCATG	3360
QY	3361	TGCGCACGCTGCAGGTTATCAAGCTCCAGAGCTGATGTGAMCAGGCGCTGGCCCTTGAGT	3420
Db	3361	TGCGCACGCTGCAGGTTATCAAGCTCCAGAGCTGATGTGAMCAGGCGCTGGCCCTTGAGT	3420
QY	3421	CATCCGAAACCTTGAACAGACAGTGAGACAAATGGAAAGCATCTCCCTTAGACAGTTCCAGCA	3480
Db	3421	CATCCGAAACCTTGAACAGACAGTGAGACAAATGGAAAGCATCTCCCTTAGACAGTTCCAGCA	3480
QY	3481	CTGCAGATGGGACACACAGACAGTGAACCATTTGACAGCCACAGACAGTAAAGCCCATGCGAG	3540
Db	3481	CTGCAGATGGGACACACAGACAGTGAACCATTTGACAGCCACAGACAGTAAAGCCCATGCGAG	3540
QY	3541	CTGTACGGCACTCAACAGGTCAAGAGAGAGGCGGCTTCTGCTCAAGAAAGAGAGCTT	3600
Db	3541	CTGTACGGCACTCAACAGGTCAAGAGAGAGGCGGCTTCTGCTCAAGAAAGAGAGCTT	3600
QY	3601	CGACACTACCTTAATAATAGTTCCAGGCCAGGAAACATATGGGGAAAGAACCCAGGAAAGATG	3660
Db	3601	CGACACTACCTTAATAATAGTTCCAGGCCAGGAAACATATGGGGAAAGAACCCAGGAAAGATG	3660
QY	3661	TTCTTGAACCTACACACAGCAAGAGCTTACTGCTGCAGCGTGCCTGTCTGGCAAGACATG	3720
Db	3661	TTCTTGAACCTACACACAGCAAGAGCTTACTGCTGCAGCGTGCCTGTCTGGCAAGACATG	3720
QY	3721	AGGTGGGTCAAGAGGGTGTAGGTGATCTGGTTGGATGAGAGAAAAGTCAAGAGAACACAG	3780
Db	3721	AGGTGGGTCAAGAGGGTGTAGGTGATCTGGTTGGATGAGAGAAAAGTCAAGAGAACACAG	3780
QY	3781	AGGCTTTTGTACACTCTTGAGACCAACAGTCAAAAGGCTGTGATGTGACATATGACAGTG	3840
Db	3781	AGGCTTTTGTACACTCTTGAGACCAACAGTCAAAAGGCTGTGATGTGACATATGACAGTG	3840
QY	3841	AAAGTATGGAGTGGCGCGGCTGTCCAGAAAAGGAGATCTGAAGTGCAGACTCTTACCC	3900
Db	3841	AAAGTATGGAGTGGCGCGGCTGTCCAGAAAAGGAGATCTGAAGTGCAGACTCTTACCC	3900
QY	3901	TGAGAGAGGAGAGATGAGAACTGACCTTGAAGAAAGAGAAAGGAGACAAAGCCACAGC	3960
Db	3901	TGAGAGAGGAGAGATGAGAACTGACCTTGAAGAAAGAGAAAGGAGACAAAGCCACAGC	3960
QY	3961	AAAGTATGGAGAAAGGTGAGACAGGAAACAGCCGCTCTGACATGAAAGAACCTTACGGGAA	4020
Db	3961	AAAGTATGGAGAAAGGTGAGACAGGAAACAGCCGCTCTGACATGAAAGAACCTTACGGGAA	4020
QY	4021	AGCCAGTCTTACACTTGACATGCTCCAGCTCAGAGAGGGGGAAAGGACATGCGGAGACCTTGG	4080
Db	4021	AGCCAGTCTTACACTTGACATGCTCCAGCTCAGAGAGGGGGAAAGGACATGCGGAGACCTTGG	4080
QY	4081	GAGGAAGCCCTTCTCTCCAGACCAGACAAAGCAGGTGCTATGAGTTCAAGTTCAAA	4140
Db	4081	GAGGAAGCCCTTCTCTCCAGACCAGACAAAGCAGGTGCTATGAGTTCAAGTTCAAA	4140
QY	4141	GCTTGGACACAAACAGTCACTTCAACACAGAAACCTGTGGAAAAGGTCAATGAAACGGTTG	4200
Db	4141	GCTTGGACACAAACAGTCACTTCAACACAGAAACCTGTGGAAAAGGTCAATGAAACGGTTG	4200
QY	4201	TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGACACTTAATTCACAGCTGAGA	4260
Db	4201	TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGACACTTAATTCACAGCTGAGA	4260
QY	4261	AGTCTCTTGCACACGGGTGGCCACTGGACTTTGACAGATGACAGAGCACGAGTACCCCTTGG	4320

Db	4261	AGTCTCTGCAAGGGATGGCACTGGACTCTTCAAGCATGCAAGGACACGGTACCCTTGG	4320
Qy	4321	GGCCTGAGTCTCAGGACGAATCCATCTCCATCATAGTAACTCTGTCTTGAAAGCACCC	4380
Db	4321	GGCCTGAGTCTCAGGACGAATCCATCTCCATCATAGTAACTCTGTCTCTGAAAGCACCC	4380
Qy	4381	TGACTCTCGACCTTCAAGGAGAAATTAAGCGCATCTCCAGAGAGCGGATCAGAGAAAGG	4440
Db	4381	TGACTCTCGACCTTCAAGGAGAAATTAAGCGCATCTCCAGAGAGCGGATCAGAGAAAGG	4440
Qy	4441	ACAAGCCAGATGGCGGTCCGATGCTGACGGGCAAGGAGAGTACAGCAATGGA AAAAGTCC	4500
Db	4441	ACAAGCCAGATGGCGGTCCGATGCTGACGGGCAAGGAGAGTACAGCAATGGA AAAAGTCC	4500
Qy	4501	TCAAGGCTGAACTTGAGATCTCTGAACTTGAGAGTAAAGNACAAAGATTGTGCTGACG	4560
Db	4501	TCAAGGCTGAACTTGAGATCTCTGAACTTGAGAGTAAAGNACAAAGATTGTGCTGAAAG	4560
Qy	4561	TCATTTCAAGACAGCCGTTTGACCAAGTTGCGACGTACAGAAACAGCCCCGAAACTATGCTT	4620
Db	4561	TCATTTCAAGACAGCCGTTTGACCAAGTTGCGACGTACAGAAACAGCCCCGAAACTATGCTT	4620
Qy	4621	ATGATTCACAGACCCAGGTTCCGTGATGCAAGGCTTGACAGAGGAGGCCAACAATGCT	4680
Db	4621	ATGATTCACAGACCCAGGTTCCGTGATGCAAGGCTTGACAGAGGAGGCCAACAATGCT	4680
Qy	4681	GGACAAAAATGAAAGATGCCAAGATGAAACACCCAGTGGCCGACGCCAGAGAGNACTTGC	4740
Db	4681	GGACAAAAATGAAAGATGCCAAGATGAAACACCCAGTGGCCGACGCCAGAGAGNACTTGC	4740
Qy	4741	AAATCTCGACCGTTCTGAGAGGCATGGGCTTAGCTTGCAAAATGCTTGGCCGCTTGACAT	4800
Db	4741	AAATCTCGACCGTTCTGAGAGGCATGGGCTTAGCTTGCAAAATGCTTGGCCGCTTGACAT	4800
Qy	4801	TGAAAAGCGCGGGTGTCAAAGTAAAGCATTTGAGAAAGCGCTCTCAACCCAAAGATCCAAA	4860
Db	4801	TGAAAAGCGCGGGTGTCAAAGTAAAGCATTTGAGAAAGCGCTCTCTCAACCCAAAGATCCAAA	4860
Qy	4861	AGGAGCATGCTGCTGATGAGCGCTCACTCCAAAGCTTTAGCCAGGACAGAGCCAGTGGCT	4920
Db	4861	AGGAGCATGCTGCTGATGAGCGCTCACTCCAAAGCTTTAGCCAGGACAGAGCCAGTGGCT	4920
Qy	4921	CTGGAACCTTAAACCAAGAAATCCCCAGACACCAAGGACCAAGCTAACCGAGAGGGCG	4980
Db	4921	CTGGAACCTTAAACCAAGAAATCCCCAGACACCAAGGACCAAGCTAACCGAGAGGGCG	4980
Qy	4981	ATCCCCCAAAGGTGAGGTTCCAGGAAGAAATGTCTACCAAGTCAGTCAAGAGAAACA	5040
Db	4981	ATCCCCCAAAGGTGAGGTTCCAGGAAGAAATGTCTACCAAGTCAGTCAAGAGAAACA	5040
Qy	5041	AGGCCCAGGACGAAGAGACTGACAGAGCCAAAGGAGACCTGGCAGAACTTAAAGATG	5100
Db	5041	AGGCCCAGGACGAAGAGACTGCTCAGAGGCAAAAGGAGACCTGGCAGAACTTAAAGATG	5100
Qy	5101	TTAGTTGCTCATTTGATCATCTGTAAAGCAGAAATGTGA AAAACAGTCAAGAAACAAGATG	5160
Db	5101	TTAGTTGCTCATTTGATCATCTGTAAAGCAGAAATGTGA AAAACAGTCAAGAAACAAGATG	5160
Qy	5161	CTGCGTTGGGACCTTGAAGCCAAAGATTTCAAGGCCCAAGABAATCCAGAGAGCAGGGCG	5220
Db	5161	CTGCGTTGGGACCTTGAAGCCAAAGATTTCAAGGCCCAAGABAATCCAGAGAGCAGGGCG	5220
Qy	5221	TCCAATGATTTCCACCAGTAGAGAGACCCGCAATTTGAGAGCTTCAATCGGAGACTAGA	5280
Db	5221	TCCAATGATTTCCACCAGTAGAGAGACCCGCAATTTGAGAGCTTCAATCGGAGACTAGA	5280
Qy	5281	GCACGCTAACATTTCTCTGTTTCAAGATGCTTGAATTTGGCCCTTGAATGCGCTCGTG	5340
Db	5281	GCACGCTAACATTTCTCTGTTTCAAGATGCTTGAATTTGGCCCTTGAATGCGCTCGTG	5340
Qy	5341	TATTTCTAACATTTCTCTGTTTCAAGATGCTTGAATTTGGCCCTTGAATGCGCTCGTG	5400
Db	5341	TATTTCTAACATTTCTCTGTTTCAAGATGCTTGAATTTGGCCCTTGAATGCGCTCGTG	5400

QY 601 GGGGAGAAAGCCCTGTGCGAGCTGGAGACCAACAGAGCCCGAGTGTGAGACTGCGCTG 660
 Db 601 GGGGAGAAAGCCCTGTGCGAGCTGGAGACCAACAGAGCCCGAGTGTGAGACTGCGCTG 660
 QY 661 GAGAGTCAGACTCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAAAGAGAGGACACC 720
 Db 661 GAGAGTCAGACTCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAAAGAGAGGACACC 720
 QY 721 TGAAGCAAGAAACAGAGCGACACAGAAATCCCTCTTCAAGCCGAAATCTGATCAAGGGCTG 780
 Db 721 TGAAGCAAGAAACAGAGCGACACAGAAATCCCTCTTCAAGCCGAAATCTGATCAAGGGCTG 780
 QY 781 AGGAAGAAAGCCAAAGATGAAGAGAGAAAGAAACAGAGAAAGAGCCCAAGGTCCCGAG 840
 Db 781 AGGAAGAAAGCCAAAGATGAAGAGAGAAAGAAACAGAGAAAGAGCCCAAGGTCCCGAG 840
 QY 841 AATCCCCAGAGCCCAAGTCAACAGTGAAGCAACATCTTCTTCAAGAAATCTTCACTC 900
 Db 841 AATCCCCAGAGCCCAAGTCAACAGTGAAGCAACATCTTCTTCAAGAAATCTTCACTC 900
 QY 901 ACGGTTGGGCGCGCTGGCGCAAGAACCAAGCTTCAAGAAATCAAAAAGAGATGATCTGG 960
 Db 901 ACGGTTGGGCGCGCTGGCGCAAGAACCAAGCTTCAAGAAATCAAAAAGAGATGATCTGG 960
 QY 961 AAACCTGCCGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 Db 961 AAACCTGCCGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 QY 1021 AGACAGAGCCAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 Db 1021 AGACAGAGCCAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 QY 1081 CAGCAGACTACGAGAAAGGTGAGAGCTGCTTGGAAAGACAGGTTGATGATGAGAGAG 1140
 Db 1081 CAGCAGACTACGAGAAAGGTGAGAGCTGCTTGGAAAGACAGGTTGATGATGAGAGAG 1140
 QY 1141 CGTCAGAGAGAGAGTGTGCTCTTGGCAAGAGTGTGATGAGAGAGAGAGAGAGAG 1200
 Db 1141 CGTCAGAGAGAGAGTGTGCTCTTGGCAAGAGTGTGATGAGAGAGAGAGAGAGAG 1200
 QY 1201 ACCAAGAAAGTGTGTCAGAGAGTCCAGTGAAGCAGCTGAGAGAGAGAGAGAGAG 1260
 Db 1201 ACCAAGAAAGTGTGTCAGAGAGTCCAGTGAAGCAGCTGAGAGAGAGAGAGAGAG 1260
 QY 1261 GAG 1320
 Db 1261 GAG 1320
 QY 1321 CTGAAGAACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 Db 1321 CTGAAGAACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 QY 1381 TGAAG 1440
 Db 1381 TGAAG 1440
 QY 1441 CTGAAG 1500
 Db 1441 CTGAAG 1500
 QY 1501 CCTCTCAGAGAAAGATCAAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 Db 1501 CCTCTCAGAGAAAGATCAAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 QY 1561 GCTTAAAGAGAGCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 Db 1561 GCTTAAAGAGAGCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 QY 1621 AGCCTGAGAGAAATCAACACATTCACACCGAAATCCCAAGAGAGAGAGAGAGAGAG 1680
 Db 1621 AGCCTGAGAGAAATCAACACATTCACACCGAAATCCCAAGAGAGAGAGAGAGAGAG 1680
 QY 1681 GAGAGAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1740

Db 1681 GAGAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1740
 QY 1741 CGCTGGAAG 1800
 Db 1741 CGCTGGAAG 1800
 QY 1801 AG 1860
 Db 1801 AG 1860
 QY 1861 TCCGAG 1920
 Db 1861 TCCGAG 1920
 QY 1921 TGTCTCTCAGTATGAG 1980
 Db 1921 TGTCTCTCAGTATGAG 1980
 QY 1981 AACAAAGCCGAG 2040
 Db 1981 AACAAAGCCGAG 2040
 QY 2041 TTTGTGTGATCATCCAG 2100
 Db 2041 TTTGTGTGATCATCCAG 2100
 QY 2101 GGGCAG 2160
 Db 2101 GGGCAG 2160
 QY 2161 CGGAG 2220
 Db 2161 CGGAG 2220
 QY 2221 CACCCGAG 2280
 Db 2221 CACCCGAG 2280
 QY 2281 AAAGATTGATCATCCAG 2340
 Db 2281 AAAGATTGATCATCCAG 2340
 QY 2341 CTAGGTGAG 2400
 Db 2341 CTAGGTGAG 2400
 QY 2401 TTAAGAAATTCATCCCGAG 2460
 Db 2401 TTAAGAAATTCATCCCGAG 2460
 QY 2461 CTGTGAG 2520
 Db 2461 CTGTGAG 2520
 QY 2521 TGCCTCTGTGAGATATATCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
 Db 2521 TGCCTCTGTGAGATATATCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
 QY 2581 AGCTGCCCAAGCTGTGAG 2640
 Db 2581 AGCTGCCCAAGCTGTGAG 2640
 QY 2641 ACATGTGAGTGTGAG 2700
 Db 2641 ACATGTGAGTGTGAG 2700
 QY 2701 CTCTCTGTGAGATATCCGCTTCCGTAACAGAACTCTTGAACACACACAGGAGAGAGAG 2760
 Db 2701 CTCTCTGTGAGATATCCGCTTCCGTAACAGAACTCTTGAACACACACAGGAGAGAGAG 2760
 QY 2761 TGCCACCTGTGAG 2820

Db 2761 TGCCACCTGTTGAAGAGGTCACTGAAAAGACATCATTTGCAGAGAAACTCCTGTGCTCA 2820
Qy 2821 CCCAGAGCTTACCAAGAGGTAAAGATGCCCATGACGATGGTCCAGTGAAATGTGATT 2880
Db 2821 CCCAGAGCTTACCAAGAGGTAAAGATGCCCATGACGATGGTCCAGTGAAATGTGATT 2880
Qy 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCCGCTACTGAAGATT 2940
Db 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCCGCTACTGAAGATT 2940
Qy 2941 CCGAAGCATCGGGGGCCGAAGAGACCAAGACATGGTGTCCGCACTGACTG 3000
Db 2941 CCGAAGCATCGGGGGCCGAAGAGACCAAGACATGGTGTCCGCACTGACTG 3000
Qy 3001 ACTCCCAGACACACAGAGGAAGCCACCCAGTTCAGGAGGTAGAGTGTGCTAG 3060
Db 3001 ACTCCCAGACACACAGAGGAAGCCACCCAGTTCAGGAGGTAGAGTGTGCTAG 3060
Qy 3061 ATACAGAAGAGAGGAGCGCCAGACGAGGCCATCTCCAAAGCCGTTCGACAGAGTGA 3120
Db 3061 ATACAGAAGAGAGGAGCGCCAGACGAGGCCATCTCCAAAGCCGTTCGACAGAGTGA 3120
Qy 3121 AAGAGGAGTCCCAGGTGCTGCAACCCAGACTGTGCGAGAGAACGGGTCAAAAGCACTGG 3180
Db 3121 AAGAGGAGTCCCAGGTGCTGCAACCCAGACTGTGCGAGAGAACGGGTCAAAAGCACTGG 3180
Qy 3181 AGAAGGTTGAGGAGGTAGAGGAGACTCCGAAGTGTGGTTCGGAGAAAGAGAGACG 3240
Db 3181 AGAAGGTTGAGGAGGTAGAGGAGACTCCGAAGTGTGGTTCGGAGAAAGAGAGACG 3240
Qy 3241 TTATGCCGAAGGACCCGTGCGAGAGCTGGAGCTGAGCATCTTGCACAGGGCTCTGAGA 3300
Db 3241 TTATGCCGAAGGACCCGTGCGAGAGCTGGAGCTGAGCATCTTGCACAGGGCTCTGAGA 3300
Qy 3301 CTGGACAGGCTACTCCAGAGAGCTTGAAGTTCCTGAAGTCAAGGAGATGTAGACCATG 3360
Db 3301 CTGGACAGGCTACTCCAGAGAGCTTGAAGTTCCTGAAGTCAAGGAGATGTAGACCATG 3360
Qy 3361 TCGCCACGTCGCGGTTATCAAGTCCAGCAGCTGATGGAAACAGGCGGTGCGCCCTGAGT 3420
Db 3361 TCGCCACGTCGCGGTTATCAAGTCCAGCAGCTGATGGAAACAGGCGGTGCGCCCTGAGT 3420
Qy 3421 CATCCGAACCTTGACACAGCTGAGCAAAATGGAAGCATCCCTTAGCAGATTTCAGACA 3480
Db 3421 CATCCGAACCTTGACACAGCTGAGCAAAATGGAAGCATCCCTTAGCAGATTTCAGACA 3480
Qy 3481 CTGCAGATGGGACACAGCAAGATGAACCTTGAACAGCAGTAAAGCCACTGCGAG 3540
Db 3481 CTGCAGATGGGACACAGCAAGATGAACCTTGAACAGCAGTAAAGCCACTGCGAG 3540
Qy 3541 CTGTCAGGCGATCAGAGTCAAGAGAGAGGCGGCTACTGTCTCAGAAAGAGAGCCTT 3600
Db 3541 CTGTCAGGCGATCAGAGTCAAGAGAGAGGCGGCTACTGTCTCAGAAAGAGAGCCTT 3600
Qy 3601 CGACACTACCTAATATGTTCCAGCCAGGAAGAACATGGGGAAGAACCCAGGAAGAGATG 3660
Db 3601 CGACACTACCTAATATGTTCCAGCCAGGAAGAACATGGGGAAGAACCCAGGAAGAGATG 3660
Qy 3661 TTCTTTGAACCTTACACAGCAGAGCTTACTGCTGAGCCGCTGCGGCTTCTGCAAGACTG 3720
Db 3661 TTCTTTGAACCTTACACAGCAGAGCTTACTGCTGAGCCGCTGCGGCTTCTGCAAGACTG 3720
Qy 3721 AGGTGGGTCAAGAGGGTGAGGTTGACTGGTTGGATGGAGAAAAGTCAAAAGAAAGCAGG 3780
Db 3721 AGGTGGGTCAAGAGGGTGAGGTTGACTGGTTGGATGGAGAAAAGTCAAAAGAAAGCAGG 3780
Qy 3781 AGGTGTTGTACACTTGGACCCCAAGTCAAAAGGCTGTGATGTGACATATGACAGTG 3840
Db 3781 AGGTGTTGTACACTTGGACCCCAAGTCAAAAGGCTGTGATGTGACATATGACAGTG 3840
Qy 3841 AAGTGATGGGAGTGGCGGGTGTGAGGAAAGGAGAGTACTGAAGTGCAGAGTCTTAGCC 3900
Db 3841 AAGTGATGGGAGTGGCGGGTGTGAGGAAAGGAGAGTACTGAAGTGCAGAGTCTTAGCC 3900

Qy 3901 TGGAGGAGGAGAGATGGAAAACCTGACGTTGAAAAGGAGAAAAGGAGACAAAAGCCAGAGC 3960
Db 3901 TGGAGGAGGAGAGATGGAAAACCTGACGTTGAAAAGGAGAAAAGGAGACAAAAGCCAGAGC 3960
Qy 3961 AAGTGAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTGAGCATGAGAGAACCTTACCGGA 4020
Db 3961 AAGTGAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTGAGCATGAGAGAACCTTACCGGA 4020
Qy 4021 AGCCAGTCTTGACACTTGACATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAGCCCTTG 4080
Db 4021 AGCCAGTCTTGACACTTGACATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAGCCCTTG 4080
Qy 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGAGGTTGCATAGAGTTCAAGTTCAAA 4140
Db 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGAGGTTGCATAGAGTTCAAGTTCAAA 4140
Qy 4141 GCCTGGACACAAAGTCACTCAAAACAGCAGCAAGCTGTGGAAAAGTTCATAGAACGGTTG 4200
Db 4141 GCCTGGACACAAAGTCACTCAAAACAGCAGCAAGCTGTGGAAAAGTTCATAGAACGGTTG 4200
Qy 4201 TGATTTTCAGACAGAGGTGAAAGTCCAGAGTGTGTAGGTGCACTTATTACCAGCTGAGA 4260
Db 4201 TGATTTTCAGACAGAGGTGAAAGTCCAGAGTGTGTAGGTGCACTTATTACCAGCTGAGA 4260
Qy 4261 AGTCCTCTGCAAAGGGTGGCCACTGGACTCTTCCAGCATGAGAGGACACGGTACCCCTGG 4320
Db 4261 AGTCCTCTGCAAAGGGTGGCCACTGGACTCTTCCAGCATGAGAGGACACGGTACCCCTGG 4320
Qy 4321 GGCTGAGTCTCAGGCGAGATCCATCCCAATCATAGTAACTCTCTGCTTCCGAAAGCACCC 4380
Db 4321 GGCTGAGTCTCAGGCGAGATCCATCCCAATCATAGTAACTCTCTGCTTCCGAAAGCACCC 4380
Qy 4381 TACATCTCTGACCTTACAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
Db 4381 TACATCTCTGACCTTACAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
Qy 4441 ACAAGCCAGATGCTGTGCTGTGATGCTGACGGCAAGGAGAGTACAGCAATCGAAAAGTCC 4500
Db 4441 ACAAGCCAGATGCTGTGCTGTGATGCTGACGGCAAGGAGAGTACAGCAATCGAAAAGTCC 4500
Qy 4501 TCAAGCTGAACTGAGATCTCTGAACTTGAAGATTAAGACAAACAGATTTGTCTGAACG 4560
Db 4501 TCAAGCTGAACTGAGATCTCTGAACTTGAAGATTAAGACAAACAGATTTGTCTGAACG 4560
Qy 4561 TCATTTCAGACAGCGCTTGACCGTTCGACGTACAGAAAACAGCCCGGAACTCATGCTT 4620
Db 4561 TCATTTCAGACAGCGCTTGACCGTTCGACGTACAGAAAACAGCCCGGAACTCATGCTT 4620
Qy 4621 ATGATTTCACAGACCCAGGTTCTTGATGAGGTTGACAGAGGAGCCCAACAGATGCT 4680
Db 4621 ATGATTTCACAGACCCAGGTTCTTGATGAGGTTGACAGAGGAGCCCAACAGATGCT 4680
Qy 4681 GGAACAAAATGAAAGATGCCCAAGATGAACACCCAGTTCGCGAGCCGAGAGAGACTTGC 4740
Db 4681 GGAACAAAATGAAAGATGCCCAAGATGAACACCCAGTTCGCGAGCCGAGAGAGACTTGC 4740
Qy 4741 AAGTCTGACCGTTCTGGAGGAGTGGCTCAGCGCTCGGAAATGCTTGCAGGCTTGCAGT 4800
Db 4741 AAGTCTGACCGTTCTGGAGGAGTGGCTCAGCGCTCGGAAATGCTTGCAGGCTTGCAGT 4800
Qy 4801 TGAAGCGCCGGTGTCAAAGTAAAGCATTTGAGAAAGTGCCTCTCAACCCCAAGATCCAAA 4860
Db 4801 TGAAGCGCCGGTGTCAAAGTAAAGCATTTGAGAAAGTGCCTCTCAACCCCAAGATCCAAA 4860
Qy 4861 AGGAGCATGCTGTGATGGCCCTCAGCTCAAAAGTTCAGCCAGGAGGAGGCGGCTGCT 4920
Db 4861 AGGAGCATGCTGTGATGGCCCTCAGCTCAAAAGTTCAGCCAGGAGGAGGCGGCTGCT 4920
Qy 4921 CTGGAACCTTAACCAAGATCCCGAGACACCAAGGACCAAGCTTAAACGAGAGGGCG 4980
Db 4921 CTGGAACCTTAACCAAGATCCCGAGACACCAAGGACCAAGCTTAAACGAGAGGGCG 4980


```
QY 4981 ATCCCCAAAAGTTGAGGTCCAGAGAAAGAAATGCTTACCAAGTCAGTCAAAAGAAC 5040
D 4981 ATCCCCAAAAGTTGAGGTCCAGAGAAAGAAATGCTTACCAAGTCAGTCAAAAGAAC 5040
QY 5041 AGGCCAGGAGAGAGAGACCTGACAGAGCCAAAGAGAGACCTGACAGATCTTAAGATG 5100
D 5041 AGGCCAGGAGAGAGAGACCTGACAGAGCCAAAGAGAGACCTGACAGATCTTAAGATG 5100
QY 5101 TTAGTTGCTCATTTGATCATCTGTGAAGACAGATGTGAAAACAGATCAAGAACAGATG 5160
D 5101 TTAGTTGCTCATTTGATCATCTGTGAAGACAGATGTGAAAACAGATCAAGAACAGATG 5160
QY 5161 CTGCTGTTGGAGCTTGGAGACCAATTTCCAGAGCCCTGATCCAGAGAGAGAGAGAG 5220
D 5161 CTGCTGTTGGAGCTTGGAGACCAATTTCCAGAGCCCTGATCCAGAGAGAGAGAGAG 5220
QY 5221 TCCAAATGATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5280
D 5221 TCCAAATGATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5280
QY 5281 GCCAGCTAATCTTCTGCTTCAAGAGCTGCTTGAATTTGCTGATGCTGCTGCTGCTG 5340
D 5281 GCCAGCTAATCTTCTGCTTCAAGAGCTGCTTGAATTTGCTGATGCTGCTGCTGCTG 5340
QY 5341 TATTTCTAATCTTCTGCTTCAAGAGCTGCTTGAATTTGCTGATGCTGCTGCTGCTG 5400
D 5341 TATTTCTAATCTTCTGCTTCAAGAGCTGCTTGAATTTGCTGATGCTGCTGCTGCTG 5400
QY 5401 TATTTGAGATTTAAGGCTGCTGCTTCAAGAGCTGCTTGAATTTGCTGATGCTGCTGCTG 5460
D 5401 TATTTGAGATTTAAGGCTGCTGCTTCAAGAGCTGCTTGAATTTGCTGATGCTGCTGCTG 5460
QY 5461 ACTTCTCAATCTGAGAGCTCTCTTATGATTTATATATATATATATATATATATATAT 5520
D 5461 ACTTCTCAATCTGAGAGCTCTCTTATGATTTATATATATATATATATATATATATAT 5520
QY 5521 TCTGTACCTATTTGATATATTTTCTTAAAGCTTGAAGCAATGCTTTGATTTATGCAA 5580
D 5521 TCTGTACCTATTTGATATATTTTCTTAAAGCTTGAAGCAATGCTTTGATTTATGCAA 5580
QY 5581 TATATTAAGGAGTGTGAGAGCATAGAGAGCTTGAAGAGCTCAAGCTCAAGCTGTAAC 5640
D 5581 TATATTAAGGAGTGTGAGAGCATAGAGAGCTTGAAGAGCTCAAGCTCAAGCTGTAAC 5640
QY 5641 TGCAGCAACAGATTAATCTCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5700
D 5641 TGCAGCAACAGATTAATCTCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5700
QY 5701 GCTTATGATCTGTGGCTTCTAGTCTCTGAAAGTGTGTTTCTTATGCAAGAGAGCT 5760
D 5701 GCTTATGATCTGTGGCTTCTAGTCTCTGAAAGTGTGTTTCTTATGCAAGAGAGCT 5760
QY 5761 CAGAAATTAACCCCATTTTGAACATCCAGAGAGTCCCATATTAACATGATTTTTC 5820
D 5761 CAGAAATTAACCCCATTTTGAACATCCAGAGAGTCCCATATTAACATGATTTTTC 5820
QY 5821 CCCCCCTTTTGTCTAATCCAGTCCAGGTGGAAGAGAGTCTCTGTGTCAATTAAGCC 5880
D 5821 CCCCCCTTTTGTCTAATCCAGTCCAGGTGGAAGAGAGTCTCTGTGTCAATTAAGCC 5880
QY 5881 CTGTCTCTTAATGATGAGCAAAATGATGTGCTTAAGGCTAGAGATGTTCTTAATGC 5940
D 5881 CTGTCTCTTAATGATGAGCAAAATGATGTGCTTAAGGCTAGAGATGTTCTTAATGC 5940
QY 5941 AGAAGAACTGTGTGAGCTTTTGTGATGTAATCTTCAAGCTGAGAGAGAGAGAGAG 6000
D 5941 AGAAGAACTGTGTGAGCTTTTGTGATGTAATCTTCAAGCTGAGAGAGAGAGAGAG 6000
QY 6001 GCAGATGAAAGTGTGCTGTTCTTCAAGATGTAATTTGATGATGATGATGATGATGATG 6060
D 6001 GCAGATGAAAGTGTGCTGTTCTTCAAGATGTAATTTGATGATGATGATGATGATGATG 6060
QY 6061 TGTGTTATATCTGTGCCCCCTTTTAAAGAAATGTTGATTAATGTTCCCTTGGATTAAT 6120
```

```
DB 6061 TGTGTTATATCTGTGCCCCCTTTTAAAGAAATGTTGATTAATGTTCCCTTGGATTAAT 6120
QY 6121 TGTGATTTGACCACTGATTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6160
D 6121 TGTGATTTGACCACTGATTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6160

RESULT 3
US-09-902-432-3
; Sequence 3, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ. ID NOS: 35
; SOFTWARE: PASCSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6160
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-3

Query Match 100.0%; Score 6160; DB 34; Length 6160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCGGGCGGAGATGAAGAGCCAGTGAAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
D 1 CGGGCGGGCGGAGATGAAGAGCCAGTGAAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 61 GAGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
D 61 GAGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 GAGCGCGAGTGAAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
D 121 GAGCGCGAGTGAAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 181 AGCTCCCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
D 181 AGCTCCCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 TCCATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
D 241 TCCATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 GACAGCAGAGAGTCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
D 301 GACAGCAGAGAGTCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 CCACAGCTGTTGAGAGATTCACAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
D 361 CCACAGCTGTTGAGAGATTCACAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 AGATCCCTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
D 421 AGATCCCTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 ATGATGTTGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
D 481 ATGATGTTGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 ATAAATAAGAAAGTCAATATCTGTCAACTACTGCTGTAAGAGAGATGAAGCGAAG 600
```


QY 2761 TGCCACTGTTGAAGAGTCACTGAAAAACATCATTTGCGAAGAAACTCTGTGTCTCA 2820
DB 2761 TGCCACTGTTGAAGAGTCACTGAAAAACATCATTTGCGAAGAAACTCTGTGTCTCA 2820
QY 2821 CCCGAGCTTAAAGAGTAAAGTCCCTGAGACATGATGACCTGAGTGAAGTGAAT 2880
DB 2821 CCCGAGCTTAAAGAGTAAAGTCCCTGAGACATGATGACCTGAGTGAAGTGAAT 2880
QY 2881 TCACCTCGAAGTGTGACAGCCACAGAGACCTCAAGAGCTCTCCGTACTGAAGATTA 2940
DB 2881 TCACCTCGAAGTGTGACAGCCACAGAGACCTCAAGAGCTCTCCGTACTGAAGATTA 2940
QY 2941 CCGAAGCATCGGGGGCCGAAGACACAGACATGTGTCCGCACTTCCAGCTGACTG 3000
DB 2941 CCGAAGCATCGGGGGCCGAAGACACAGACATGTGTCCGCACTTCCAGCTGACTG 3000
QY 3001 ACTCCCGACAGACACAGAGAGCCACCCAGTTCAAGAGGTGAGAGTGTGTCTAG 3060
DB 3001 ACTCCCGACAGACACAGAGAGCCACCCAGTTCAAGAGGTGAGAGTGTGTCTAG 3060
QY 3061 ATACAGAAAGAGAGAGCCGACAGAGCCATCTCCAAAGCCGTTGACAGCAAGTGA 3120
DB 3061 ATACAGAAAGAGAGAGCCGACAGAGCCATCTCCAAAGCCGTTGACAGCAAGTGA 3120
QY 3121 AAGAGAGTCCAGGTGCTTGCAACCTCAGACTGTGCAAGAAAGGGGTCAAAAGCATGG 3180
DB 3121 AAGAGAGTCCAGGTGCTTGCAACCTCAGACTGTGCAAGAAAGGGGTCAAAAGCATGG 3180
QY 3181 AAGAGGTGAGAGGTGAGAGAGAGCTCCGAAGTGTGCTTCCGAGAAAGAGAGAGCG 3240
DB 3181 AAGAGGTGAGAGGTGAGAGAGAGCTCCGAAGTGTGCTTCCGAGAAAGAGAGAGCG 3240
QY 3241 TTATGCCGAAAGAGCCCGTGACAGAGCTGAGACTGAGCATCTTGCAGAGGGCTCTGAGA 3300
DB 3241 TTATGCCGAAAGAGCCCGTGACAGAGCTGAGACTGAGCATCTTGCAGAGGGCTCTGAGA 3300
QY 3301 CTGACAGAGCTTCTCAGAGAGCTTGAAGTCTCTGAAGTCAAGCAAGATGTAAGCATG 3360
DB 3301 CTGACAGAGCTTCTCAGAGAGCTTGAAGTCTCTGAAGTCAAGCAAGATGTAAGCATG 3360
QY 3361 TCGCAGAGTGCAGGTATTCAGAGCTCCAGAGCTGATGGAAGAGCCGAGCCCTGAGT 3420
DB 3361 TCGCAGAGTGCAGGTATTCAGAGCTCCAGAGCTGATGGAAGAGCCGAGCCCTGAGT 3420
QY 3421 CATCCGAAGCTTGCAGACAGTGAAGAAATGGAAGCATCTCTTACAGAGATTCAGACA 3480
DB 3421 CATCCGAAGCTTGCAGACAGTGAAGAAATGGAAGCATCTCTTACAGAGATTCAGACA 3480
QY 3481 CTGCAGATGGGACAGAGCAAGATGAACCATTTGACAGCCAGGACAGTAAAGCCACTGCA 3540
DB 3481 CTGCAGATGGGACAGAGCAAGATGAACCATTTGACAGCCAGGACAGTAAAGCCACTGCA 3540
QY 3541 CTGTAGGAGTCAAGGTCAAGAGAAAGAGGCGGCTTCTGCTCAAGAAAGAGAGCCCT 3600
DB 3541 CTGTAGGAGTCAAGGTCAAGAGAAAGAGGCGGCTTCTGCTCAAGAAAGAGAGCCCT 3600
QY 3601 CGACACTACTTAATATGTTCCAGCCAGAGAAACATGAGGAGAAACCCAGAGAGATG 3660
DB 3601 CGACACTACTTAATATGTTCCAGCCAGAGAAACATGAGGAGAAACCCAGAGAGATG 3660
QY 3661 TTCTTGAACCTTACAGAGCAAGAGCTTACTGCTGACAGCCGTGCTTCTGCGAAAGACTG 3720
DB 3661 TTCTTGAACCTTACAGAGCAAGAGCTTACTGCTGACAGCCGTGCTTCTGCGAAAGACTG 3720
QY 3721 AGGTGGGTCAAGAGGTGAGTGTGATGTGAGAGAAAGTCAAGAAAGAAAGCAAG 3780
DB 3721 AGGTGGGTCAAGAGGTGAGTGTGATGTGAGAGAAAGTCAAGAAAGAAAGCAAG 3780
QY 3781 AGGTGTTGTATCACTGTGAGCCCAAGCATGCAAAAGGCTGATGTGACATTTGACAGTG 3840
DB 3781 AGGTGTTGTATCACTGTGAGCCCAAGCATGCAAAAGGCTGATGTGACATTTGACAGTG 3840

QY 3841 AAGTATGGAGTGGCCGGGTGTCAAGAAAAAGAGATCACTGAAGTGCAGAGTCTTACG 3900
DB 3841 AAGTATGGAGTGGCCGGGTGTCAAGAAAAAGAGATCACTGAAGTGCAGAGTCTTACG 3900
QY 3901 TGGAGAGGAGATGGAATACTGACGTTGAAAAAGAGAAAGGAGACAAAGCCAGAGC 3960
DB 3901 TGGAGAGGAGATGGAATACTGACGTTGAAAAAGAGAAAGGAGACAAAGCCAGAGC 3960
QY 3961 AAGTATGGAAGAGGTGAGACAGAAACAGCCGCTCTTGAGCATGGAAGAACTTACGGGA 4020
DB 3961 AAGTATGGAAGAGGTGAGACAGAAACAGCCGCTCTTGAGCATGGAAGAACTTACGGGA 4020
QY 4021 AGCCAGTCTGACACTTGAATGCCAGCTCAGAGAGGGGAAAGGCACTGAGGAAGCTTG 4080
DB 4021 AGCCAGTCTGACACTTGAATGCCAGCTCAGAGAGGGGAAAGGCACTGAGGAAGCTTG 4080
QY 4081 GAGGAGCCCTTCTCTCCAGACCAAGACAAAGCAGTTGCATAGAGTTCAAGTTCAAA 4140
DB 4081 GAGGAGCCCTTCTCTCCAGACCAAGACAAAGCAGTTGCATAGAGTTCAAGTTCAAA 4140
QY 4141 GCTTGAACAAACAGTCACTTCAACAGACAGAGCTGTGAAAAAGTATGAAAAAGTGG 4200
DB 4141 GCTTGAACAAACAGTCACTTCAACAGACAGAGCTGTGAAAAAGTATGAAAAAGTGG 4200
QY 4201 TGATTTGAGACAGGTGAAGTCCAGAGTGTGAGTGACACTTATTAACAAGCTGAGA 4260
DB 4201 TGATTTGAGACAGGTGAAGTCCAGAGTGTGAGTGACACTTATTAACAAGCTGAGA 4260
QY 4261 AGTCTCTGCAACGGGTGGCCACTGGACTCTTGCAGTGCAGAGACACGGTACCCCTGG 4320
DB 4261 AGTCTCTGCAACGGGTGGCCACTGGACTCTTGCAGTGCAGAGACACGGTACCCCTGG 4320
QY 4321 GGCTGAGTCTCAGGCAAGATCCATCCCAATCATAGTAACTCTGCTCTGAAAGCATCC 4380
DB 4321 GGCTGAGTCTCAGGCAAGATCCATCCCAATCATAGTAACTCTGCTCTGAAAGCATCC 4380
QY 4381 TACATCTGACCTTCAAGAGGAATTAAGCGATCCGAGAGAGCGATTCAGAGGAAGG 4440
DB 4381 TACATCTGACCTTCAAGAGGAATTAAGCGATCCGAGAGAGCGATTCAGAGGAAGG 4440
QY 4441 ACAAGCCAGATGCTGTCTCTGATCTGACGCAAGAGAGAGTACAGCAATCGAAAAATGCC 4500
DB 4441 ACAAGCCAGATGCTGTCTCTGATCTGACGCAAGAGAGAGTACAGCAATCGAAAAATGCC 4500
QY 4501 TCAAGGCTGAACCTGAGATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4560
DB 4501 TCAAGGCTGAACCTGAGATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4560
QY 4561 TCATTGAGACGCGTTGACCAAGTTCGACGTAAGAAACAGCCCGGAAACATGATGCTT 4620
DB 4561 TCATTGAGACGCGTTGACCAAGTTCGACGTAAGAAACAGCCCGGAAACATGATGCTT 4620
QY 4621 ATGATTCAGAACCCAGTTCTCTGATCTGACGCAAGGCTTGAACAGAGGAGCCCAACAGATCT 4680
DB 4621 ATGATTCAGAACCCAGTTCTCTGATCTGACGCAAGGCTTGAACAGAGGAGCCCAACAGATCT 4680
QY 4681 GGAACAAAATGAAGATGCAAGATGAAGAACCAAGTGCAGCCGACCAAGAGAGACTTGC 4740
DB 4681 GGAACAAAATGAAGATGCAAGATGAAGAACCAAGTGCAGCCGACCAAGAGAGACTTGC 4740
QY 4741 AAGTCTGACCGTTCTGAGAGCATGGGCTCAGCTCGGAAATGCTTGGCGGCTTGCAGT 4800
DB 4741 AAGTCTGACCGTTCTGAGAGCATGGGCTCAGCTCGGAAATGCTTGGCGGCTTGCAGT 4800
QY 4801 TGAAGGCGCGGTGTCAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 4860
DB 4801 TGAAGGCGCGGTGTCAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 4860
QY 4861 AAGAGCATGTCTGATGAGCCCTCAGCTCAAGCTTAAGCCAGGAGAGGCAAGTGCCT 4920
DB 4861 AAGAGCATGTCTGATGAGCCCTCAGCTCAAGCTTAAGCCAGGAGAGGCAAGTGCCT 4920
QY 4921 CTGGAACCTTACCAAGATATCCCAAGACCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4980

Db 4921 CTGAAACCTTAACAAAGAAATCCCAAGACCAACGACCAAGCTAACCGAGGAGCGG 4980
Qy ATCCCCCAAAAGTTGAGTCCAGGAAGAAATGTTCTACCAAGTCAAGTCAAGAGAAACA 5040
Db ATCCCCCAAAAGTTGAGTCCAGGAAGAAATGTTCTACCAAGTCAAGTCAAGAGAAACA 5040
Qy AGGCCAGGCAAG 5100
Db AGGCCAGGCAAG 5100
Qy TTAGTTGCTCATTTACATCTGTAAAGACCAAGAAATGTGAAGAAACAGTCCACAGAAAGATG 5160
Db TTAGTTGCTCATTTACATCTGTAAAGACCAAGAAATGTGAAGAAACAGTCCACAGAAAGATG 5160
Qy CTGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGAGAGAGAG 5220
Db CTGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGAGAGAGAG 5220
Qy TCCAAATGATTTCCACACAGTAGAGACACCCGACAAATTTCTGAGGCTTCATCGGAGCTAGA 5280
Db TCCAAATGATTTCCACACAGTAGAGACACCCGACAAATTTCTGAGGCTTCATCGGAGCTAGA 5280
Qy GCAGCTAAACATTTCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGATGCGCTCCGCTG 5340
Db GCAGCTAAACATTTCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGATGCGCTCCGCTG 5340
Qy TATTTCTAAACATTTCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGATGCGCTCCGCTG 5400
Db TATTTCTAAACATTTCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGATGCGCTCCGCTG 5400
Qy TATTTCCGATTTAAGTCTCGGTTCTCAACCTGGAAACCAATTCATGCGATACCTAGTTCC 5460
Db TATTTCCGATTTAAGTCTCGGTTCTCAACCTGGAAACCAATTCATGCGATACCTAGTTCC 5460
Qy ACTTCTCAAACTGAGAGATCTCTCTTTATGTTATGTTATGTTATGTTATGTTATGTTATG 5520
Db ACTTCTCAAACTGAGAGATCTCTCTTTATGTTATGTTATGTTATGTTATGTTATGTTATG 5520
Qy TCCTGTACCTATTGTTATATTTTCTAAAGTCTTAAAGCAATGCTTTTATGTTATGTTATG 5580
Db TCCTGTACCTATTGTTATATATTTTCTAAAGTCTTAAAGCAATGCTTTTATGTTATGTTATG 5580
Qy TATATAAGGCTGTGAGCAGATAGCGAGCTTTGAAAGCTTCAAGCTCAACCTGTTAACT 5640
Db TATATAAGGCTGTGAGCAGATAGCGAGCTTTGAAAGCTTCAAGCTCAACCTGTTAACT 5640
Qy TGCAAGCAACAGATAACATTTCTGCGAAGAGAGACAGAGCTTTTAAAGTTTACTGAT 5700
Db TGCAAGCAACAGATAACATTTCTGCGAAGAGAGACAGAGCTTTTAAAGTTTACTGAT 5700
Qy GCTTAGATCTGTGGGCTTTCTAGTCTCTGAAAGTGGTTGTTTCTATGACAGCGAGCT 5760
Db GCTTAGATCTGTGGGCTTTCTAGTCTCTGAAAGTGGTTGTTTCTATGACAGCGAGCT 5760
Qy CAGAAATATAAAACCCCAATTTGAAACATCCAGGATGTCCTCAATATACCATGATTTTTC 5820
Db CAGAAATATAAAACCCCAATTTGAAACATCCAGGATGTCCTCAATATACCATGATTTTTC 5820
Qy CCCCCCTTTTGTCTAATCCAGTCCAGGTTGGAAGAGAGTCTCTCTGTGTGAGATTAAGCC 5880
Db CCCCCCTTTTGTCTAATCCAGTCCAGGTTGGAAGAGAGTCTCTCTGTGTGAGATTAAGCC 5880
Qy CTGTCTCTAATGATATGGAACAAATGATGTCCTAAGGCCATGAGATGTTTCTCTAATGC 5940
Db CTGTCTCTAATGATATGGAACAAATGATGTCCTAAGGCCATGAGATGTTTCTCTAATGC 5940
Qy AGAAGGAATCTGTTGACGTTTTTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 6000
Db AGAAGGAATCTGTTGACGTTTTTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 6000
Qy GCAGATCGAAGTGAATCTGTTTCTTTACAGATGTTATTTGATGATGATGATGATGATGATG 6060

Db 6001 GCAGATCGAAGTGAATCTGTTTCTTTACAGATGTTATTTGATAGATGATGATGATG 6060
Qy TGTGTTATATCTGTGCCCCCTTTTAAAGAAACAATGTTGCAATATGTTCTTTCGATAAAT 6120
Db TGTGTTATATCTGTGCCCCCTTTTAAAGAAACAATGTTGCAATATGTTCTTTCGATAAAT 6120
Qy TGTGATTTGCAACTGATTTAAATAAACAATATTTGACTAC 6160
Db TGTGATTTGCAACTGATTTAAATAAACAATATTTGACTAC 6160

RESULT 4

PCT-US02-16173A-2063
; Sequence 2063, Application PC/TUS0216173A
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-WO
; CURRENT APPLICATION NUMBER: PCT/US02/16173A
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2063
; LENGTH: 5236
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_057103
PCT-US02-16173A-2063

Query Match 84.5%; Score 5204.2; DB 1; Length 5236;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 5223; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

Qy 1 CGGGCCGGGGAGTAGAAGAGCCACTGAGCCATGGGCGCAGGAGTTCACCGAGCAGC 60
Db 1 CGGGCCGGGGAGTAGAAGAGCCACTGAGCCATGGGCGCAGGAGTTCACCGAGCAGC 60
Qy 61 GGAGCCCCGAGAGCCGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 61 GGAGCCCCGAGAGCCGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Qy 121 GGCCCCGAGCTGAAGCTCGGGAGAGCTGGAGACCCCGCCGAGCGGAGCCCGCCACCA 180
Db 121 GGCCCCGAGCTGAAGCTCGGGAGAGCTGGAGACCCCGCCGAGCGGAGCCCGCCACCA 180
Qy 181 AGCTCCCAAGAAATGGCCAGCTGTCTTCTCAACGCGCTAGCTGTCTTCTCAACAGAGATG 240

Db	181	AGCTCCACAGAGAAATGGCCACCTGTCTTCTGTCAAGCGCTACCTGAACAGAGATG	240
Qy	241	TCCATGTCCAAAGAGAAAAACCAAGAGGGGCGAGAGAAAGTCTGTANTGAGATGTTG	300
Db	241	TCCTATGTCCAAGAGGAAACCAAGAGGGGCGAGAGAAAGTCTGTATGAGATGTTG	300
Qy	301	GACACGCCAGAGTCAAGAGATGTGTAGAGAAAAAGATCCGAGTTGAAGAAATTTGCGGCTCAACT	360
Db	301	GACAGCCAGAGTCAAGAGATGTGTAGAGAAAAAGATCCGAGTTGAAGAAATTTGCGGCTCAACT	360
Qy	361	CCACAGCGTTGAAGATATCAAAAGATGGGCGAGAGAGACATCAGAAATTAATTGTAAC	420
Db	361	CCACAGCGTTGAAGATATCAAAAGATGGGCGAGAGAGACATCAGAAATTAATTGTAAC	420
Qy	421	AGATCCCTGCTTCAGAAAACAATGTGGAGAAATGTACAGCTCTGTAGTCCAGGCTA	480
Db	421	AGATCCCTGCTTCAGAAAACAATGTGGAGAAATGTATACGCTCTGTAGTCCAGGCTA	480
Qy	481	ATGATGTTGGCTTCAGAAAGATTAATTAATTTGTTGTTTAAATTCACGGTGAAGAG	540
Db	481	ATGATGTTGGCTTCAGAAAGATTAATTAATTTGTTGTTTAAATTCACGGTGAAGAG	540
Qy	541	ATAAAAATGAAATTCAGATATCTGTCCAACTACTCATCTCAAGAAAGATGAAGCGCAG	600
Db	541	ATAAAAATGAAATTCAGATATCTGTCCAACTACTCATCTCAAGAAAGATGAAGCGCAG	600
Qy	601	GGGCGAAGCTCTGTGTGGAGCTGGAGACCAACAGAGGCCCATGTGTGAGATCGCGTGG	660
Db	601	GGGCGAAGCTCTGTGTGGAGCTGGAGACCAACAGAGGCCCATGTGTGAGATCGCGTGG	660
Qy	661	GAGAGTCAGATCCAAAGAAATGTAGCTGAAGCAATCCAAGAGAACAGAGAGCACCC	720
Db	661	GAGAGTCAGATCCAAAGAAATGTAGCTGAAGCAATCCAAGAGAAAGCAAGAGCACCC	720
Qy	721	TGAAGCAAGAACAGAGACGACAGAAATTCCTCCCTTAAACCCGATCTGATCAAGCGGTG	780
Db	721	TGAAGCAAGAACAGAGACGACAGAAATTCCTCCCTTAAACCCGATCTGATCAAGCGGTG	780
Qy	781	AGGAAAGAACCAAGAGATGAAAGGAGAGAAAGAAACAGAGAAAGGCCACCAAGTCCCGAG	840
Db	781	AGGAAAGAACCAAGAGATGAAAGGAGAGAGAAAGAAACAGAGAAAGGCCACCAAGTCCCGAG	840
Qy	841	AATCCCCGAGACGCCCACTCAACAGTGTAGACCAACATCTTCTTCAAGAAATTCTTCACTC	900
Db	841	AATCCCCGAGACGCCCACTCAACAGTGTAGACCAACATCTTCTTCAAGAAATTCTTCACTC	900
Qy	901	ACGGTTGGGCGGCTGGCGCAGAGACCAAGCTTCAAGAAATCMAAAGAGATGATCTGG	960
Db	901	ACGGTTGGGCGGCTGGCGCAGAGACCAAGCTTCAAGAAATCMAAAGAGATGATCTGG	960
Qy	961	AAACCTGCCGAAABAAGAAAGGACGCAAGGCAAAAAATGACGAGAGAAAGAAAGSAA	1020
Db	961	AAACCTGCCGAAABAAGAAAGGACGCAAGGCAAAAAATGACGAGAGAAAGAAAGSAA	1020
Qy	1021	AGACAGAGCCAGCTCTCGAGAGACGAGAGCGCGCAAGACACAGACAGGCGCAGTTGT	1080
Db	1021	AGACAGAGCCAGCTCTCGAGAGAGCGAGAGCGCGCAAGACACAGACAGGCGCAGTTGT	1080
Qy	1081	CAGCAGACTACGAGAGGTGAGCTGCTTTGGCAACGGAATGTTTATGTAGAGAGAGAAAGCC	1140
Db	1081	CAGCAGACTACGAGAGGTGAGCTGCTTTGGCAACGAGGTTGGTGAACCTTGGAGGAT	1140
Qy	1141	CGTCAGAGAGAAAGTGTCTCTTTGGCAACGGAATGTTTATGTAGAGAGAGAAAGCC	1200
Db	1141	CGTCAGAGAGAAAGTGTGTCTTTTGGCAACGGAAGTGTGTATGTAGAGAGAGAAAGCC	1200
Qy	1201	ACCAAGAGTTGTTGCAAGAGGTCCACTGTAGACCCGTGAGAGAAACAGAGAGAGAGCAG	1260
Db	1201	ACCAAGAGTTGTTGCAAGAGGTCCACTGTAGACCCGTGAGAGAAACAGAGAGAGAGCAG	1260
Qy	1261	GAGGAGAGAGAGAGCTGAAGGGGCGTGTGTAGAGAGAACAGAGATCTTGGCCCC	1320

Db	1261	GAGGAGGAGGAGGCTGAAAGGAGGAGCGTGTGTGTGAAGAGAAACAAGGAATCTTTGCCCC	1322
Qy	1321	CTGGAACACTGGCTGAGCCCCAGAGAGTCCCCAGAGAGCTGTGAGTGAAGACTGA	1380
Db	1321	CTGGAACACTGGCTGAGCCCCAGAGAGTCCCCAGAGAGCTGTGAGTGAAGACTGA	1380
Qy	1381	TGAAGACAGAGAGATGTGTGTCTCTGAGAGAGACCAACTCACTGACAGACTTAAGTC	1440
Db	1381	TGAAGACAGAGAGATGTGTGTCTCTGAGAGAGACCAACTCACTGAGACTTAAGTC	1440
Qy	1441	CTGAAGAGAGAGCGCTGCCCAAAACCCAGAGAGGCAATTGTCAGTGAGGTGAGATGCTGT	1500
Db	1441	CTGAAGAGAGAGCGCTGCCCAAAACCCAGAGAGGCAATTGTCAGTGAGGTGAGATGCTGT	1500
Qy	1501	CCTCTCAGGAAAGATCAGATGATCAGAGGAATCTCTTGAAGAACTCTTCAGTGTCTAG	1560
Db	1501	CCTCTCAGGAAAGATCAGATGATCAGAGGAATCTCTTGAAGAACTCTTCAGTGTCTAG	1560
Qy	1561	GCTTAAAGAACTGTCTGAGGAGAGACCAAGGGGAAACGAGAGGTGTGGGAGAACAGAG	1620
Db	1561	GCTTAAAGAACTGTCTGAGGAGAGACCAAGGGGAAACGAGAGGTGTGGGAGAACAGAG	1620
Qy	1621	AGCTCGAGAAATACCAACATTCACACCGAATCCCGAGAGTGTGATGACAGAGG	1680
Db	1621	AGCTCGAGAAATACCAACATTCACACCGAATCCCGAGAGTGTGATGACAGAGAG	1680
Qy	1681	GAGAGAGCTGTGCGTGTGCTCCCGAGAGAGCTTGAGAGAACCACTGTCTGAGAAAGGC	1740
Db	1681	GAGAGAGCTGTGCGTGTGCTCCCGAGAGAGCTTGAGAGAACCACTGTGTGAGAAAGGC	1740
Qy	1741	CGCTGAGAGACCCCAAGATGGGAGAGCTGAGAGAGAACTAATTCCGATGAGAGAGAA	1800
Db	1741	CGCTGAGAGACCCCAAGATGGGAGAGCTGAGAGAGAACTAATTCCGATGAGAGAGAA	1800
Qy	1801	AGAGAGAAAGGATATCACTCCCTGGGCATCTCTTAAAAAGATGGTGACACCCAGAGAAACGG	1860
Db	1801	AGAGAGAAAGGATATCACTCCCTGGGCATCTCTTAAAAAGATGGTGACACCCAGAGAAACGG	1860
Qy	1861	TCCGAGAAGCTTCTGAGAGTGAACAGAGAGAAAGCTGAGAGAGTCAAGAGCCCACT	1920
Db	1861	TCCGAGAAGCTTCTGAGAGTGAACAGAGAGAAAGCTGAGAGAGTCAAGAGCCCACT	1920
Qy	1921	TGTCTCTCACTGATAGCACAGTGTCAAGAAATGCAAGTGAAGTCAAACTGTGTGTGAG	1980
Db	1921	TGTCTCTCACTGATAGCACAGTGTCAAGAAATGCAAGTGAAGTCAAACTGTGTGTGAG	1980
Qy	1981	AACAAAAGCCAGAGAGAACCAAGACGTAGAGGTGATTACTTCAGTGTCTTGGAGACACTGA	2040
Db	1981	AACAAAAGCCAGAGAGAACCAAGACGTAGAGGTGATTACTTCAGTGTCTTGGAGACACTGA	2040
Qy	2041	TTTGTGTGCGATCATCCAGAGAGAGAGCAAGAGGCAATCTCTTCAGATGATGAAGAG	2100
Db	2041	TTTGTGTGCGATCATCCAGAGAGAGAGCAAGAGGCAATCTCTTCAGATGATGAAGAG	2100
Qy	2101	GCGCAAGACACTGGGAGGGGACATGACAGAGACAGAGAGGCGCAGCAAGACAAAGAG	2160
Db	2101	GCGCAAGACACTGGGAGGGGACATGACAGAGACAGAGAGGCGCAGCAAGAGAAAG	2160
Qy	2161	CCGGAACAGAGCGTGTCTTCGACAGACCCAGAGAGAGAGCAACACGCAAGAAAGTTCT	2220
Db	2161	CCGGAACAGAGCGTGTCTTCGACAGACCCAGAGAGAGAGCAACACGCAAGAAAGTTCT	2220
Qy	2221	CACCCGAGCCAGCGGAGAGGCTTCTCGAAGGGGAGAGTGTCTCCACTTGGGAGTCAATTA	2280
Db	2221	CACCCGAGCCAGCGGAGAGGCTTCTCGAAGGGGAGAGTGTCTCCACTTGGGAGTCAATTA	2280
Qy	2281	AAAGATTAGTCACTCCAGAGAAAAATCCAGTCAAAACTGAGAGAGAAAGCCAGAGCT	2340
Db	2281	AAAGATTAGTCACTCCAGAGAAAAATCCAGTCAAAACTGAGAGAGAAAGCCAGAGCT	2340
Qy	2341	CTAGTGTAGAGCAATGTGTCCACTGAGATGGAACCGAGTGTGAGAAAGAAATCTTGGGTTTCA	2400
Db	2341	CTAGTGTAGAGCAATGTGTGTCCACTGAGATGGAACCGAGTGTGAGAAAGAAATCTTGGGTTTCA	2400

QY 2401 TTAAGAAATTCATCCCGGACGGCGGAAGAAAAGGGCGACGCGGAAGCAAGCA 2460
DB 2401 TTAAGAAATTCATCCCGGACGGCGGAAGAAAAGGGCGAGATGGGAAGCAAGCA 2460
QY 2461 CTGTGGAAGACTCAGGGCCAGTGGAGATAAATAGGACGACCTTAATGTCTCCACCGCTCG 2520
DB 2461 CTGTGGAAGACTCAGGGCCAGTGGAGATAAATAGGACGACCTTAATGTCTCCACCGCTCG 2520
QY 2521 TGCCTCTGTCTGAGTATAATGACGTGGAGGGGAGAGATGGAAGCCACAGGGGAATACGG 2580
DB 2521 TGCCTCTGTCTGAGTATGATGACGTGGAGGGGAGAGATGGAAGCCACAGGGGAATGCGG 2580
QY 2581 AGCTGCCCCAGCTCTCGGGGCTGTGTACGTGTCGAGGAGCTCAGTAAGACTCTGTGTC 2640
DB 2581 AGCTGCCCCAGCTCTCGGGGCTGTGTACGTGTCGAGGAGCTCAGTAAGACTCTGTGTC 2640
QY 2641 ACACCTGTGAGTGTGCGAGTCAATTCATGGGACACGGGAGTCAACAGTGTGCGAAGCGGT 2700
DB 2641 ACACCTGTGAGTGTGCGAGTCAATTCATGGGACACGGGAGTCAACAGTGTGCGAAGCGGT 2700
QY 2701 CTCCTTCGTGGATATCCGCTTCCGTAAACAGAACCTCTTTGAACACACAGCGGGAGGCCA 2760
DB 2701 CTCCTTCGTGGATATCCGCTTCCGTAAACAGAACCTCTTTGAACACACAGCGGGAGGCCA 2760
QY 2761 TGCACCTGTTGAAGAGTCTACTGAAAAGACATCATTTGCAAGAAACTCTCTGTGCTCA 2820
DB 2761 TGCACCTGTTGAAGAGTCTACTGAAAAGACATCATTTGCAAGAAACTCTCTGTGCTCA 2820
QY 2821 CCCAGAGCTTACCAGAGGGTAAAGATGCCATGACGATGCTCACCAGTGAAGTGATT 2880
DB 2821 CCCAGAGCTTACCAGAGGGTAAAGATGCCATGACGATGCTCACCAGTGAAGTGATT 2880
QY 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAGTTA 2940
DB 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAGTTA 2940
QY 2941 CCGAAGCATCGGGGGCCGAAGAGACCAAGACATGGTTCGCGAGTTTCCAGCTGACTG 3000
DB 2941 CCGAAGCATCGGGGGCCGAAGAGACCAAGACATGGTTCGCGAGTTTCCAGCTGACTG 3000
QY 3001 ACTCCACAGACACACAGAGGAAGCCACCCAGTTCCAGGAGTAGAGTGGTGTAG 3060
DB 3001 ACTCCACAGACACACAGAGGAAGCCACCCAGTTCCAGGAGTAGAGTGGTGTAG 3060
QY 3061 ATACAGAAGAGAGGAGCGCCAGACGAGCCATCTCCAAAGCCGTTGCAGACAAGGTGA 3120
DB 3061 ATACAGAAGAGAGGAGCGCCAGACGAGCCATCTCCAAAGCCGTTGCAGACAAGGTGA 3120
QY 3121 AAGAGGATCCAGGTGCTGCAACCCAGACTGTGCAGAGAACGGGTTCAAAAGCACTGG 3180
DB 3121 AAGAGGATCCAGGTGCTGCAACCCAGACTGTGCAGAGAACGGGTTCAAAAGCACTGG 3180
QY 3181 AGAAGGTTGAGGAGGTAGAGAGACTCCGAAAGTGTGGCTTCGGGAGAAAGAGAGACG 3240
DB 3181 AGAAGGTTGAGGAGGTAGAGAGACTCCGAAAGTGTGGCTTCGGGAGAAAGAGAGACG 3240
QY 3241 TTATGCCGAAAGGACCCGCTGCAGGAAGCTGGAGCTGAGCATCTTGCACAGGGCTCTGAGA 3300
DB 3241 TTATGCCGAAAGGACCCGCTGCAGGAAGCTGGAGCTGAGCATCTTGCACAGGGCTCTGAGA 3300
QY 3301 CTGACAGGCTTACTCCAGAGAGCCCTTGAAGTTCTGAAGTTCACGCGCAGATGTAGACCATG 3360
DB 3301 CTGACAGGCTTACTCCAGAGAGCCCTTGAAGTTCACGCGCAGATGTAGACCATG 3360
QY 3361 TCGCCAGCTGCGAGGTTATCAAGCTCCAGAGCTGATGGAACAGGCGCTGCGCCCTGAGT 3420
DB 3361 TCGCCAGCTGCGAGGTTATCAAGCTCCAGAGCTGATGGAACAGGCGCTGCGCCCTGAGT 3420
QY 3421 CATCCGAAACCTTCACAGACAGTCAAGCAATGGAACACTCCCTTAGCAGATTTCAGACA 3480
DB 3421 CATCCGAAACCTTCACAGACAGTCAAGCAATGGAACACTCCCTTAGCAGATTTCAGACA 3480

QY 3481 CTCAGATGGGACACAGCAAGATGAAACCAATTCACAGCCAGGACAGTAAGCCACTGCAG 3540
DB 3481 CTCAGATGGGACACACAGCAAGATGAAACCAATTCACAGCCAGGACAGTAAGCCACTGCAG 3540
QY 3541 CTGTCAAGGAGTCAACAGTTCACAGAGAGAGAGCGGCTACTGTCTCAGAAAGAGAGCCTT 3600
DB 3541 CTGTCAAGGAGTCAACAGTTCACAGAGAGAGAGCGGCTACTGTCTCAGAAAGAGAGCCTT 3600
QY 3601 CGACACTACCTAATAATGTTTCCAGCCCCAGGAAGAACATGCGGGAAGAGATG 3660
DB 3601 CGACACTACCTAATAATGTTTCCAGCCCCAGGAAGAACATGCGGGAAGAGATG 3660
QY 3661 TTCTTGAACCTTACACAGCAAGAGCTTACTGCTCAGCCGCTGCCCTTCTGCGAAAGACTG 3720
DB 3661 TTCTTGAACCTTACACAGCAAGAGCTTACTGCTCAGCCGCTGCCCTTCTGCGAAAGACTG 3720
QY 3721 AGTGGGTCAAGAGGGTGAAGTTGACTGGTTGGATGAGAGAAAAGTCAAAAGAAAGACAGG 3780
DB 3721 AGTGGGTCAAGAGGGTGAAGTTGACTGGTTGGATGAGAGAAAAGTCAAAAGAAAGACAGG 3780
QY 3781 AGTGTGTTTGTACACTCTCTGGACCCCAACAGTCAAAAGGCTCTGATGTGACATATGACAGTG 3840
DB 3781 AGTGTGTTTGTACACTCTCTGGACCCCAACAGTCAAAAGGCTCTGATGTGACATATGACAGTG 3840
QY 3841 AAGTGTGGAGTGGCGGGTGTCTCAGAAAAGGAGAGTACTGAAGTGCAGAGCTTTAGCC 3900
DB 3841 AAGTGTGGAGTGGCGGGTGTCTCAGAAAAGGAGAGTACTGAAGTGCAGAGCTTTAGCC 3900
QY 3901 TGGAGGAGGAGAGATGGAACCTGAGTTGAAAGAGAAAAGGGAGACAAAGCCAGAGC 3960
DB 3901 TGGAGGAGGAGAGATGGAACCTGAGTTGAAAGAGAAAAGGGAGACAAAGCCAGAGC 3960
QY 3961 AAGTGTGAAAGAGGTGAGCAGGAAAACAGCCGCTCTCTGAGCATGAAGAAACCTACCGGA 4020
DB 3961 AAGTGTGAAAGAGGTGAGCAGGAAAACAGCCGCTCTCTGAGCATGAAGAAACCTACCGGA 4020
QY 4021 AGCCAGTCTTGAACCTTCCAGCATGCCAGTCTCAGAGGGGGAAAGGCACCTGGGAAGCCTTG 4080
DB 4021 AGCCAGTCTTGAACCTTCCAGCATGCCAGTCTCAGAGGGGGAAAGGCACCTGGGAAGCCTTG 4080
QY 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGGTTCAAGTTCAAA 4140
DB 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGCATAGAGGTTCAAGTTCAAA 4140
QY 4141 GCCTTGCACACAACAGTCACTCAAAACAGCAGAAAGCTGTGGAAAAGGTCATAGAAACGGTTG 4200
DB 4141 GCCTTGCACACAACAGTCACTCAAAACAGCAGAAAGCTGTGGAAAAGGTCATAGAAACGGTTG 4200
QY 4201 TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGTGCACACTTATTAACAGCTGAGA 4260
DB 4201 TGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGTGCACACTTATTAACAGCTGAGA 4260
QY 4261 AGTCCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGGACACGGTACCCCTGG 4320
DB 4261 AGTCCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGGACACGGTACCCCTGG 4320
QY 4321 GGCTTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACTCTCTGCTCTTGAAGACACC 4380
DB 4321 GGCTTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACTCTCTGCTCTTGAAGACACC 4380
QY 4381 TACATCTCTGACCTTACAGGAGAAATAAGCGCATCCCGAGAGAGCGGATCAGAGGAAGAGG 4440
DB 4381 TACATCTCTGACCTTACAGGAGAAATAAGCGCATCCCGAGAGAGCGGATCAGAGGAAGAGG 4440
QY 4441 ACAAGCCAGATGCTGGTCTCTGATGCTGACGGCAAGAGAGTACAGCAATCGAAAAAGTCC 4500
DB 4441 ACAAGCCAGATGCTGGTCTCTGATGCTGACGGCAAGAGAGTACAGCAATCGAAAAAGTCC 4500
QY 4501 TCAAGGCTGAACCTGAGATCTCGAACTTTCAGAGTGAAGCAACAAGATTGCTGNAACG 4560
DB 4501 TCAAGGCTGAACCTGAGATCTCGAACTTTCAGAGTGAAGCAACAAGATTGCTGNAACG 4560
QY 4561 TCATTTCAGACAGCCGTTGACCACTGTCGACGTACAGAAAACAGCCCCCGGAACTCATGCTT 4620

```
|||||
Db 4561 TCATTGACAGCGCGTTGACCGATTGCACTGACGAAACAGCCCCCAAACTCATGCTT 4620
Qy 4621 ATGATTCACAGACCCAGGTTCTTCATGACAGGCTTGACAGAGGCCAAACAGATCT 4680
Db 4621 ATGATTCACAGACCCAGGTTCTTCATGACAGGCTTGACAGAGGCCAAACAGATCT 4680
Qy 4681 GGACAAAATGAAAGATGCCAAGATGAAACACCCAGTCCGACGCCAGAGAGGATCTGC 4740
Db 4681 GGACAAAATGAAAGATGCCAAGATGAAACACCCAGTCCGACGCCAGAGAGGATCTGC 4740
Qy 4741 AAGTCCTGACCGTTCTGAGGATGGGCTCAGCTCGGAAATGTCTGCGCGCTTGAGT 4800
Db 4741 AAGTCCTGACCGTTCTGAGGATGGGCTCAGCTCGGAAATGTCTGCGCGCTTGAGT 4800
Qy 4801 TGAAGCCCGGCTGTCAAAGTAAGCATGAGAGCTGCTCTCAACCCAAAGATCCAA 4860
Db 4801 TGAAGCCCGGCTGTCAAAGTAAGCATGAGAGCTGCTCTCAACCCAAAGAT -CAA 4859
Qy 4861 AGAGCATGCTGCTGATGAGCCCTCAGCTCCAAAGCTTACGCCAGAGGCGCAGTGCCT 4920
Db 4861 AGAGCATGCTGCTGATGAGCCCTCAGCTCCAAAGCTTACGCCAGAGGCGCAGTGCCT 4919
Qy 4921 CTGGAACCTTAACCAAAATCCCAACACCAAGGCAAAAGCTTACGCCAGAGGCGG 4980
Db 4920 CTGGAACCTTAACCAAAAGATCCCAACACCAAGGCAAAAGCTTACGCCAGAGGCGG 4979
Qy 4981 ATCCCCAAAGTTGAGGCTCCAGAGAAAGATGTCTACCAAGTCACTCAAGAGAA 5040
Db 4980 ATGCCCAAAAGTTGAGGCTCCAGAGAAAGATGTCTACCAAGTCACTCAAGAGAA 5039
Qy 5041 AGGCCAGGCGAGAGAGAGCTGCGAGAGCCAAAGGAGAGCTTGCAGATCTTAAGTG 5100
Db 5040 AGGCCAGGCGAGAGAGAGCTGCGAGAGCCAAAGGAGAGCTTGCAGATCTTAAGTG 5099
Qy 5101 TTAGTGTCTCATTTGACATCTGTAAAGACCAAGATGTGAAAACAAATCCAGAA 5160
Db 5100 TTAGTGTCTCATTTGACATCTGTAAAGACCAAGATGTGAAAACAAAGATG 5159
Qy 5161 CTGCTGTGGGACCTTGAGACCAAGATTTCAAGGCCATGATCCAGAGAGGCGCG 5220
Db 5160 CTGCTGTGGGACCTTGAGACCAAGATTTCAAGGCCATGATCCAGAGAGGCGCG 5219
Qy 5221 TCCAAATGATTTCCACC 5237
Db 5220 TCCAAATGATTTCCACC 5236

RESULT 5
US-10-191-803-279
; Sequence 279, Application US/10191803
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5098US
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 279
```

```
LENGTH: 5236
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_057103
US-10-191-803-279

Query Match      84.5%; Score 5204.2; DB 41; Length 5236;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 5223; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

Qy 1 CGGCGCGCGGAGTAAAGACCACTGAGCCATGGCGGAGGAGTTCACCGAGAGC 60
Db 1 CGGCGCGCGGAGTAAAGACCACTGAGCCATGGCGGAGGAGTTCACCGAGAGC 60
Qy 61 GGAGCCCCGAGAGCCGCGGAGGAGCAACGCCGAGAGCTGCTTCAATGAGGATG 120
Db 61 GGAGCCCCGAGAGCCGCGGAGGAGCAACGCCGAGAGCTGCTTCAATGAGGATG 120
Qy 121 GGCGCCGAGCTGAAGCTTCGGGAGCACTGAGACCCCGCGACCGGACCCGCCA 180
Db 121 GGCGCCGAGCTGAAGCTTCGGGAGCACTGAGACCCCGCGACCGGACCCGCCA 180
Qy 181 AGCTCCCAAGAAATGCGCAAGCTCTTCTGTCACACGCGTGAACAAGAGATG 240
Db 181 AGCTCCCAAGAAATGCGCAAGCTCTTCTGTCACACGCGTGAACAAGAGATG 240
Qy 241 TCCATGTCGAAGAGAAACCAAGAGGAGGAGAGAGAAAGTGTGATGAGATGTTG 300
Db 241 TCCATGTCGAAGAGAAACCAAGAGGAGGAGAGAGAAAGTGTGATGAGATGTTG 300
Qy 301 GACAGCGAGATCGAAGATGTGAGAAAAAGCCAGTTGAAGAAATGGCGGCACT 360
Db 301 GACAGCGAGATCGAAGATGTGAGAAAAAGCCAGTTGAAGAAATGGCGGCACT 360
Qy 361 CCACAGCTGTGAAGATATCAAAAGATGGGCGAGAGAGACATCAGAAATATTAAT 420
Db 361 CCACAGCTGTGAAGATATCAAAAGATGGGCGAGAGAGACATCAGAAATATTAAT 420
Qy 421 AGATCCCTGCTTCAAGAAACCAATGTGAAAGATGTGACAGCTGCTGAGTCCAG 480
Db 421 AGATCCCTGCTTCAAGAAACCAATGTGAAAGATGTGACAGCTGCTGAGTCCAG 480
Qy 481 ATGATGTTGGCTTCAAGAAATATTTAATTTGTTGTTTAAATTCACGGTGAAG 540
Db 481 ATGATGTTGGCTTCAAGAAATATTTAATTTGTTGTTTAAATTCACGGTGAAG 540
Qy 541 ATAAATATGAAAGTCAAGATCTGTCCACTACCTGTCAGAAAGATGAAGGCAAG 600
Db 541 ATAAATATGAAAGTCAAGATCTGTCCACTACCTGTCAGAAAGATGAAGGCAAG 600
Qy 601 GGGGAGAAAGCTTGTGAGAGCTGAGAGCAACAGAGGCCAGTGTGAGACTGCGT 660
Db 601 GGGGAGAAAGCTTGTGAGAGCTGAGAGCAACAGAGGCCAGTGTGAGACTGCGT 660
Qy 661 GAGAGTCAGCATCCAAAGAAAGTGAAGCTGAAGCAATCCACAGAAAGCAAGAGC 720
Db 661 GAGAGTCAGCATCCAAAGAAAGTGAAGCTGAAGCAATCCACAGAAAGCAAGAGC 720
Qy 721 TGAAGCAAGAGCAGAGCAGCAAGAAATCCCTTCAAGCCGAATCTGATCAAGCG 780
Db 721 TGAAGCAAGAGCAGAGCAGCAAGAAATCCCTTCAAGCCGAATCTGATCAAGCG 780
Qy 781 AGGAAGAAAGCCAAAGATGAAGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAG 840
Db 781 AGGAAGAAAGCCAAAGATGAAGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAG 840
Qy 841 AATCCCGAGAGAGCCAGCTCAACAGTGAACAATCTTCTTCAAGAGTTCTTCACT 900
Db 841 AATCCCGAGAGAGCCAGCTCAACAGTGAACAATCTTCTTCAAGAGTTCTTCACT 900
Qy 901 ACGTTGGCGCGCTGGGAGAGAGACAGCTTCAAGAAATCAAGAGAGATGATCGG 960
Db 901 ACGTTGGCGCGCTGGGAGAGAGACAGCTTCAAGAAATCAAGAGAGATGATCGG 960
```


Db 901 ACGGTTGGCCGGCTGGCGCAAGAACACCAGCTTCAAGAAATCAAAAGAGGATGATCTGG 960
Qy 961 AAATGCCGAGAGAGAAAGAGCAAGAGGAGAGAAAGTAGACGAGAGAAAGAGAAA 1020
Db 961 AAATGCCGAGAGAGAAAGAGCAAGAGGAGAGAAAGTAGACGAGAGAAAGAGAAA 1020
Qy 1021 AGACAGAGCCAGCTCCGAGGAGCAGGAGCCGAGAGACACAGAGACACAGCCAGGTTGT 1080
Db 1021 AGACAGAGCCAGCTCCGAGGAGCAGGAGCCGAGAGACACAGAGACACAGCCAGGTTGT 1080
Qy 1081 CAGCAGACTACGAGAAGGTGAGCTGCCCTTTGGAGACACAGTTTGGTGCACCTGGAGGCAT 1140
Db 1081 CAGCAGACTACGAGAAGGTGAGCTGCCCTTTGGAGACACAGTTTGGTGCACCTGGAGGCAT 1140
Qy 1141 CGTCAGAGGAGAGTGTCTCTTTGGCAACGGAAGTGTTTGATGAGAAGATGGAAGCCC 1200
Db 1141 CGTCAGAGGAGAGTGTCTCTTTGGCAACGGAAGTGTTTGATGAGAAGATGGAAGCCC 1200
Qy 1201 ACCAAGAAGTTGTTCAGAGGTCCACGCTGACACCGTGGAGAACACAGAGGAGCAGG 1260
Db 1201 ACCAAGAAGTTGTTCAGAGGTCCACGCTGACACCGTGGAGAACACAGAGGAGCAGG 1260
Qy 1261 GAGGAGGAGAGAGGCTGAAGGGGGCGTGTGTGTAGAGGAACAGGAGAAATCTTTGCCCC 1320
Db 1261 GAGGAGGAGAGAGGCTGAAGGGGGCGTGTGTGTAGAGGAACAGGAGAAATCTTTGCCCC 1320
Qy 1321 CTGAGAAATCTGGCTGAGCCCCCAGAGGTCTCCCAAGGAAGCTGAGCCTGCTGAGAGCTGA 1380
Db 1321 CTGAGAAATCTGGCTGAGCCCCCAGAGGTCTCCCAAGGAAGCTGAGCCTGCTGAGAGCTGA 1380
Qy 1381 TGAAGACAGAGAGATGTGTCTCTGGAGGAGACCACTCAACTGACAGACCTAAGTC 1440
Db 1381 TGAAGACAGAGAGATGTGTCTCTGGAGGAGACCACTCAACTGACAGACCTAAGTC 1440
Qy 1441 CTGAAGAGAACGCTGCCAAAACCCAGAAAGCATTTGTCAAGTGAGGTGAGAGTCTGT 1500
Db 1441 CTGAAGAGAACGCTGCCAAAACCCAGAAAGCATTTGTCAAGTGAGGTGAGAGTCTGT 1500
Qy 1501 CCTCTCAGGAAGAAATCAAGTACAGGAACTCTTGAAGAAATCTTTCAGTAGCTCAG 1560
Db 1501 CCTCTCAGGAAGAAATCAAGTACAGGAACTCTTGAAGAAATCTTTCAGTAGCTCAG 1560
Qy 1561 GCTTAAAGAGCTCTCTGGGAAGACGAAAGGGGAAACGAGGAGGTGGGGGAGACGAAG 1620
Db 1561 GCTTAAAGAGCTCTCTGGGAAGACGAAAGGGGAAACGAGGAGGTGGGGGAGACGAAG 1620
Qy 1621 AGCCTGGAGAAATCAACACATTCACACCGAATCCCCAGAGAGTCTGATGAGCAGAGG 1680
Db 1621 AGCCTGGAGAAATCAACACATTCACACCGAATCCCCAGAGAGTCTGATGAGCAGAGG 1680
Qy 1681 GAGAGAGCTCTGCGTCTGCCCGGAGGAGCTGAGGAGACACGCTGTCTGGAGAAAGGC 1740
Db 1681 GAGAGAGCTCTGCGTCTGCCCGGAGGAGCTGAGGAGACACGCTGTCTGGAGAAAGGC 1740
Qy 1741 CGCTGGAGACCCAGGATGGGAAGCTGAGGAAGGAATCTACTTCGATGGAGAGAAGA 1800
Db 1741 CGCTGGAGACCCAGGATGGGAAGCTGAGGAAGGAATCTACTTCGATGGAGAGAAGA 1800
Qy 1801 AGACAGAGGAGTCACTCCCTGGGCATCCTTCAAAAAGATCGTGACACCCCAAGAAACGG 1860
Db 1801 AGACAGAGGAGTCACTCCCTGGGCATCCTTCAAAAAGATCGTGACACCCCAAGAAACGG 1860
Qy 1861 TCCGAAGACCTTCTGAGAGTGAAGGAGGAGAGCTGGAGAAGGTCAAGAGCGCCACCT 1920
Db 1861 TCCGAAGACCTTCTGAGAGTGAAGGAGGAGAGCTGGAGAAGGTCAAGAGCGCCACCT 1920
Qy 1921 TGTCTCCACTGATAGCAGCTGTCAAGAAATGCAAGATGAAGTCAAAACTCTTGGTAGG 1980
Db 1921 TGTCTCCACTGATAGCAGCTGTCAAGAAATGCAAGATGAAGTCAAAACTCTTGGTAGG 1980
Qy 1981 AACAAAAGCCAGAGGAACCAAGCGTAGGGTGGATACCTTCAGTCTCTGGGAAGCACTGA 2040
Db 1981 AACAAAAGCCAGAGGAACCAAGCGTAGGGTGGATACCTTCAGTCTCTGGGAAGCACTGA 2040

Qy 2041 TTTGTGTCCGATCATCCAGAAAGAGAGCAAGGAAGGCATCCTCTTCAGATGATGAAGGAG 2100
Db 2041 TTTGTGTCCGATCATCCAGAAAGAGAGCAAGGAAGGCATCCTCTTCAGATGATGAAGGAG 2100
Qy 2101 GGCCAAAGGACACTGGGAGGGGACAGTCCACAGACAGAGGAGGCCAGCAAAAGACAAAG 2160
Db 2101 GGCCAAAGGACACTGGGAGGGGACAGTCCACAGACAGAGGAGGCCAGCAAAAGACAAAG 2160
Qy 2161 CCGGAAACAGACGCTGTTCTGCCAGCACCCAGGAGCAGGACCAAGCGCAAGGAAGTTTCT 2220
Db 2161 CCGGAAACAGACGCTGTTCTGCCAGCACCCAGGAGCAGGACCAAGCGCAAGGAAGTTTCT 2220
Qy 2221 CACCCGAGCCAGCGGGAGGCCCTTCGAAAGGGAAAGTGTCTCCACTTGGGAGTCAATTTA 2280
Db 2221 CACCCGAGCCAGCGGGAGGCCCTTCGAAAGGGAAAGTGTCTCCACTTGGGAGTCAATTTA 2280
Qy 2281 AAAGATTAGTCACTCCAAAGAAAAAATCCAAAGTCAAAACTGGGAAGAGAAAGCCGAGACT 2340
Db 2281 AAAGATTAGTCACTCCAAAGAAAAAATCCAAAGTCAAAACTGGGAAGAGAAAGCCGAGACT 2340
Qy 2341 CTAGTGTAGAGCAGTTGTCACCTGAGATCGAACCGAGTAGAGAGAAATCTTTGGGTTTCCA 2400
Db 2341 CTAGTGTAGAGCAGTTGTCACCTGAGATCGAACCGAGTAGAGAGAAATCTTTGGGTTTCCA 2400
Qy 2401 TTAAGAAATTCATCCCGGACCGCGGAGAAAGGGGAGACGGGAGCAAGCAAGGCCA 2460
Db 2401 TTAAGAAATTCATCCCGGACCGCGGAGAAAGGGGAGATGGGAAGCAAGCAAGGCCA 2460
Qy 2461 CTGTGGAAGACTCAGGCGCAGTGGAGATAAATGAGGACGACCTAATGTCCAGCCCTCG 2520
Db 2461 CTGTGGAAGACTCAGGCGCAGTGGAGATAAATGAGGACGACCTAATGTCCAGCCCTCG 2520
Qy 2521 TGCTCTGTCTGAGTATAATGTCAGTGGAGAGGAGAGATGGAAGCCAGGGGAATACGG 2580
Db 2521 TGCTCTGTCTGAGTATAATGTCAGTGGAGAGGAGAGATGGAAGCCAGGGGAATACGG 2580
Qy 2581 AGTGTCCCAAGCTGTGGGGCTGTGTAACGAGAACCTCTTGAACACACAGCGGGAGAGCCA 2640
Db 2581 AGTGTCCCAAGCTGTGGGGCTGTGTAACGAGAACCTCTTGAACACACAGCGGGAGAGCCA 2640
Qy 2641 ACATGTGAGTGTGCGAGTCAATTTGATGGGACCCAGGGCAGTCAACAGTGTGGAAGCGGT 2700
Db 2641 ACATGTGAGTGTGCGAGTCAATTTGATGGGACCCAGGGCAGTCAACAGTGTGGAAGCGGT 2700
Qy 2701 CTCTCTCGTGATATCCGCTTCCGTAACAGAACCTCTTGAACACACAGCGGGAGAGCCA 2760
Db 2701 CTCTCTCGTGATATCCGCTTCCGTAACAGAACCTCTTGAACACACAGCGGGAGAGCCA 2760
Qy 2761 TGCCACCTGTTGAAGAGGTCACTGAAAAAGACATCATTTGCAGAAAGAACTCTCTGTGCTCA 2820
Db 2761 TGCCACCTGTTGAAGAGGTCACTGAAAAAGACATCATTTGCAGAAAGAACTCTCTGTGCTCA 2820
Qy 2821 CCAGAGCTTACAGAGGGTAAAGATGCCATGACGACATGGTCAACAGTGAAGTGAAT 2880
Db 2821 CCAGAGCTTACAGAGGGTAAAGATGCCATGACGACATGGTCAACAGTGAAGTGAAT 2880
Qy 2881 TCACCTCAGAAAGTGTGACAGGCCACAGAGACCTCCAGGGCTCTCCGACTGGAAGAGTTA 2940
Db 2881 TCACCTCAGAAAGTGTGACAGGCCACAGAGACCTCCAGGGCTCTCCGACTGGAAGAGTTA 2940
Qy 2941 CCGAAGCATCGGGGGCGAAGAGACCAAGATGTTGCCAGTTCACAGCTGACTG 3000
Db 2941 CCGAAGCATCGGGGGCGAAGAGACCAAGATGTTGCCAGTTCACAGCTGACTG 3000
Qy 3001 ACTCCCAGACACACAGAGAAAGCCACCCAGTTAGAGAGGTAGAGAGTGGTGTCTAG 3060
Db 3001 ACTCCCAGACACACAGAGAAAGCCACCCAGTTAGAGAGGTAGAGAGTGGTGTCTAG 3060
Qy 3061 ATACAGAAAGAGGAGCGCCAGACGAGCCATCTCCAAAGCGTTGCGAGACAAGGTGA 3120
Db 3061 ATACAGAAAGAGGAGCGCCAGACGAGCCATCTCCAAAGCGTTGCGAGACAAGGTGA 3120

QY 3121 AAGAGAGTCCAGGTGCTGCAACCCAGACTGTGTCAGAGAACGGGGTCAAAAGCACTGG 3180
 Db 3121 AAGAGAGTCCAGGTGCTGCAACCCAGACTGTGTCAGAGAACGGGGTCAAAAGCACTGG 3180
 QY 3181 AGAAGTTGAGAGGTGAGAGAGAGACTCCGAAGTGTGCTTCCGAGAGAAAGAGAGAGCG 3240
 Db 3181 AGAAGTTGAGAGGTGAGAGAGAGACTCCGAAGTGTGCTTCCGAGAGAAAGAGAGAGCG 3240
 QY 3241 TTATGCCGAAGAGACCCCGTGCAGAGAGCTGAGACTGAGCATTTTGCAAGGGCTTGAGA 3300
 Db 3241 TTATGCCGAAGAGACCCCGTGCAGAGAGCTGAGACTGAGCATTTTGCAAGGGCTTGAGA 3300
 QY 3301 CTGACAGAGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCAAGCCAGATGAGACATG 3360
 Db 3301 CTGACAGAGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCAAGCCAGATGAGACATG 3360
 QY 3361 TCGCAGCTGCGAGGTATTCAGAGCTCCAGAGCTGATGAGAACAGGCCGTGCCCCCTGAGT 3420
 Db 3361 TCGCAGCTGCGAGGTATTCAGAGCTCCAGAGCTGATGAGAACAGGCCGTGCCCCCTGAGT 3420
 QY 3421 CATCCGAACCTTGACAGACAGTGAACAATGAGACATCTCCCTTACAGATTCAGACA 3480
 Db 3421 CATCCGAACCTTGACAGACAGTGAACAATGAGACATCTCCCTTACAGATTCAGACA 3480
 QY 3481 CTGACAGATGGGACAGAGAGTGAACCATTTGACAGCCAGAGCAGTAAAGCCACTGACG 3540
 Db 3481 CTGACAGATGGGACAGAGAGTGAACCATTTGACAGCCAGAGCAGTAAAGCCACTGACG 3540
 QY 3541 CTGTCAGAGCTGACAGGTTCACAGAGAGAGAGCCGCTACTGCTCAGAGAGAGAGCCCTT 3600
 Db 3541 CTGTCAGAGCTGACAGGTTCACAGAGAGAGAGCCGCTACTGCTCAGAGAGAGAGCCCTT 3600
 QY 3601 CGACACTACTATATATGTTCACGCCACAGAGAAACATGCGGAGAAACACAGAGAAAGATG 3660
 Db 3601 CGACACTACTATATATGTTCACGCCACAGAGAAACATGCGGAGAAACACAGAGAAAGATG 3660
 QY 3661 TTCTTGAACCTACAGAGAGAGCTTACTGCTGCAAGCCGTGCGCTTCTGCGAGAAAGCTG 3720
 Db 3661 TTCTTGAACCTACAGAGAGAGCTTACTGCTGCAAGCCGTGCGCTTCTGCGAGAAAGCTG 3720
 QY 3721 AGGTGGCTCAAGAGGGGTGAGTGTGACTGTTGATGAGAGAAATTCAGAGAGAGAGAG 3780
 Db 3721 AGGTGGCTCAAGAGGGGTGAGTGTGACTGTTGATGAGAGAAATTCAGAGAGAGAGAG 3780
 QY 3781 AGGTGTTGTACACTCTGGAACCAAGTCAAAAGGCTGCTGATGTGACATGTGACAGTG 3840
 Db 3781 AGGTGTTGTACACTCTGGAACCAAGTCAAAAGGCTGCTGATGTGACATGTGACAGTG 3840
 QY 3841 AAGTATGGAGTGGCCCGGGTGTCAAGAGAAAGAGATGATGAAATGACAGATCTTAGCC 3900
 Db 3841 AAGTATGGAGTGGCCCGGGTGTCAAGAGAAAGAGATGATGAAATGACAGATCTTAGCC 3900
 QY 3901 TGGAGAGAGAGAGATGAGAACTGACGTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGC 3960
 Db 3901 TGGAGAGAGAGAGATGAGAACTGACGTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGC 3960
 QY 3961 AAGTATGGAAGAGGTGAG 4020
 Db 3961 AAGTATGGAAGAGGTGAG 4020
 QY 4021 AGCCAGTCTGACACTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4080
 Db 4021 AGCCAGTCTGACACTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4080
 QY 4081 GAG 4140
 Db 4081 GAG 4140
 QY 4141 GCTTGAAG 4200
 Db 4141 GCTTGAAG 4200
 QY 4201 TGATTTCAAG 4260

Db 4201 TGATTTCAAG 4260
 QY 4261 AGTCTCTGCAACGGGTGGCCACTGGAATCTTTCAGCATGACAGAGAGAGAGAGAGAGAG 4320
 Db 4261 AGTCTCTGCAACGGGTGGCCACTGGAATCTTTCAGCATGACAGAGAGAGAGAGAGAGAG 4320
 QY 4321 GGCCTGAGTCTCAG 4380
 Db 4321 GGCCTGAGTCTCAG 4380
 QY 4381 TACATCTGACCTTCAAG 4440
 Db 4381 TACATCTGACCTTCAAG 4440
 QY 4441 ACAAGCAGATGCTGTGCTTGCATGCTGAACGCGCAAGAGAGAGAGAGAGAGAGAGAGAG 4500
 Db 4441 ACAAGCAGATGCTGTGCTTGCATGCTGAACGCGCAAGAGAGAGAGAGAGAGAGAGAGAG 4500
 QY 4501 TCAAGGCTGAACCTGAGATCCTGGAACCTTGAAGATGAGAGAGAGAGAGAGAGAGAGAGAG 4560
 Db 4501 TCAAGGCTGAACCTGAGATCCTGGAACCTTGAAGATGAGAGAGAGAGAGAGAGAGAGAGAG 4560
 QY 4561 TCATTGAGACAGCCGTTGACAGATTCGACAGTTCGACAGAGAGAGAGAGAGAGAGAGAGAG 4620
 Db 4561 TCATTGAGACAGCCGTTGACAGATTCGACAGTTCGACAGAGAGAGAGAGAGAGAGAGAGAG 4620
 QY 4621 ATGATTCACAGACCCAGGTTCTGCAATGACAGCTTGAACAGAGAGAGAGAGAGAGAGAG 4680
 Db 4621 ATGATTCACAGACCCAGGTTCTGCAATGACAGCTTGAACAGAGAGAGAGAGAGAGAGAG 4680
 QY 4681 GGACAAAATGAAGATGACAGAGATGAAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740
 Db 4681 GGACAAAATGAAGATGACAGAGATGAAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740
 QY 4741 AAGTCTGACCGTTCTGAG 4800
 Db 4741 AAGTCTGACCGTTCTGAG 4800
 QY 4801 TGAAGCGCCGGTGTCAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 4860
 Db 4801 TGAAGCGCCGGTGTCAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 4860
 QY 4861 AGGAGCATGCTGTGATGAG 4920
 Db 4861 AGGAGCATGCTGTGATGAG 4920
 QY 4921 CTGGAACCTTAACCAAGAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980
 Db 4921 CTGGAACCTTAACCAAGAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980
 QY 4981 ATCCCCCAAGAGATGAG 5040
 Db 4981 ATCCCCCAAGAGATGAG 5040
 QY 5041 AGGCCAG 5100
 Db 5041 AGGCCAG 5100
 QY 5101 TTAAGTGTCTATTTGATCACTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5160
 Db 5101 TTAAGTGTCTATTTGATCACTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5160
 QY 5161 CTGCTGTGGAG 5220
 Db 5161 CTGCTGTGGAG 5220
 QY 5221 TCCATGATTTCCAGCC 5237
 Db 5221 TCCATGATTTCCAGCC 5237

RESULT 6

US-09-978-277A-3
; Sequence 3, Application US/08978277A
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,277A
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/665,401
; FILING DATE: 18-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
US-09-978-277A-3

Query Match 84.4%; Score 5200; DB 13; Length 5200;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 ATGGGCGCAGGAGTTCACCGAGCAGCGGAGCCCGGAGCGCGGGGAGCGCACG 92
Db 1 ATGGGCGCAGGAGTTCACCGAGCAGCGGAGCCCGGAGCGCGGGGAGCGCACG 60

Qy 93 CCGAGCAGCTGTGCTCAGTGGCCATGGGCCCGCAGCTGAAGCCTCGGAGCAGCTGGA 152
Db 61 CCGAGCAGCTGTGCTCAGTGGCCATGGGCCCGCAGCTGAAGCCTCGGAGCAGCTGGA 120

Qy 153 GACCCCGCGAGCGGACCCCGCCACCAAGCTCCACAGAGATGGCCAGCTCTTCT 212
Db 121 GACCCCGCGAGCGGACCCCGCCACCAAGCTCCACAGAGATGGCCAGCTCTTCT 180

Qy 213 GTCAACCGCTAGCTGAACAGGAGATGTCCATGTCCAAGAGGAAAAACAGGAGGGCAG 272
Db 181 GTCAACCGCTAGCTGAACAGGAGATGTCCATGTCCAAGAGGAAAAACAGGAGGGCAG 240

Qy 273 GAGGAAGAAGTCGTTGATGAGGATGTTGGACAGCGAGAGTCAAGAATGTGAGAGAAAA 332
Db 241 GAGGAAGAAGTCGTTGATGAGGATGTTGGACAGCGAGAGTCAAGAATGTGAGAGAAAA 300

Qy 333 GACCGAGTTGAAGAAATGGCGGCGCAACTCCACAGCTGTGAGATATCANAAGGATGGG 392
Db 301 GACCGAGTTGAAGAAATGGCGGCGCAACTCCACAGCTGTGAGATATCANAAGGATGGG 360

Qy 393 CAGGAGGAGACATCAGAAATAATTGAACAGATCCCTGCTTTCAGAAACAATGTGGAAGAA 452
Db 361 CAGGAGGAGACATCAGAAATAATTGAACAGATCCCTGCTTTCAGAAACAATGTGGAAGAA 420

Qy 453 ATGTTACAGCTGCTGAGTCCAGGCTTAATGATGTTGGCTTCAAGAAAGTATTAAATTT 512
Db 421 ATGTTACAGCTGCTGAGTCCAGGCTTAATGATGTTGGCTTCAAGAAAGTATTAAATTT 480

Qy 513 GTTGGTTTAAATTTCAAGGTGAAGAGGATAAAAAATGAAAGTCAGATACTGTCCAACTA 572
Db 481 GTTGGTTTAAATTTCAAGGTGAAGAGGATAAAAAATGAAAGTCAGATACTGTCCAACTA 540

Qy 573 CTCATGTTCAAGAGGATGAAGCGGAGAGGGGAGAGCCCTCTCTGTGGAGCTGGAGACCAC 632
Db 541 CTCATGTTCAAGAGGATGAAGCGGAGAGGGGAGAGCCCTCTCTGTGGAGCTGGAGACCAC 600

Qy 633 CAGGAGCCAGTGTGGAGAGCTGCGCTCGGAGAGTCAAGATCCCAAGAAAGTGAAGCTGAAG 692
Db 601 CAGGAGCCAGTGTGGAGAGCTGCGCTCGGAGAGTCAAGATCCCAAGAAAGTGAAGCTGAAG 660

Qy 693 CAATCCACAGAGAGCAAGCAAGGAGCACTTGAAGCAAGAAACAGAGCAGCAGAGAAATCCCC 752
Db 661 CAATCCACAGAGAGCAAGGAGCACTTGAAGCAAGAAACAGAGCAGCAGAGAAATCCCC 720

Qy 753 CTTCAAGCCGAATCTGATCAAGCGGCTGAGGAGAGAGCCAAAGATGAAGGAGAGAAAAA 812
Db 721 CTTCAAGCCGAATCTGATCAAGCGGCTGAGGAGAGAGCCAAAGATGAAGGAGAGAAAAA 780

Qy 813 CAAGAGAAAGAGCCCAACCAAGTCCCAGATCCCGGAGAGCCCAAGTCAACAGTGAAGACA 872
Db 781 CAAGAGAAAGAGCCCAACCAAGTCCCAGATCCCGGAGAGCCCAAGTCAACAGTGAAGACA 840

Qy 873 ACATCTTCTTCAAGAAAGTCTTCACTCAAGTGGCGCGCTGGCGCAAGAGACAGCAGC 932
Db 841 ACATCTTCTTCAAGAAAGTCTTCACTCAAGTGGCGCGCTGGCGCAAGAGACAGCAGC 900

Qy 933 TTCAAGAAATCAAAAGAGGATGATCTGGAAACTGCCGAGAGAGAAAGAGCAAGAGGCA 992
Db 901 TTCAAGAAATCAAAAGAGGATGATCTGGAAACTGCCGAGAGAGAAAGAGCAAGAGGCA 960

Qy 993 GAAAAAGTAGACGAGGAG 1052
Db 961 GAAAAAGTAGACGAGGAG 1020

Qy 1053 GCAGAGACACAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1112
Db 1021 GCAGAGACACAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080

Qy 1113 GAAGACAGGTTGGTGAACCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1172
Db 1081 GAAGACAGGTTGGTGAACCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140

Qy 1173 GAAGTGTTCATGAGAGAGATGGAAGCCCAAGAGAGTGTTCAGAGAGTTCAGAGTGAAG 1232
Db 1141 GAAGTGTTCATGAGAGAGATGGAAGCCCAAGAGAGTGTTCAGAGAGTTCAGAGTGAAG 1200

Qy 1233 ACCGTGAGAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1292
Db 1201 ACCGTGAGAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260

Qy 1293 GTAGAGAGGAAACAGGAGAAATCCTTGGCCCTTGAAGAACTGGCTGAGCCCAAGAGAGTCCCC 1352
Db 1261 GTAGAGAGGAAACAGGAGAAATCCTTGGCCCTTGAAGAACTGGCTGAGCCCAAGAGAGTCCCC 1320

Qy 1353 CAGGAAGCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAAGGA 1412
Db 1321 CAGGAAGCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAAGGA 1380

Qy 1413 GACCACACTCAACTGACAGACCTAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
Db 1381 GACCACACTCAACTGACAGACCTAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440

QY 1473 GGCAATTGACGAGGAGTGGAGATGCTGTCTCTCAGAAAGATCAGGAAAGT 1532
 Db 1441 GGCAATTGACGAGGAGTGGAGATGCTGTCTCTCAGAAAGATCAGGAAAGT 1500
 QY 1533 CCGTTGAAGAACTCTTCAAGTACGAGCTTTAAAGAACTGTCTGGAGAGAGAGAG 1592
 Db 1501 CCGTTGAAGAACTCTTCAAGTACGAGCTTTAAAGAACTGTCTGGAGAGAGAGAG 1560
 QY 1593 GGGAAACGAGAGGTGGGGGAGACGAAGACCTGGAGAAATCAACATTCACCCGA 1652
 Db 1561 GGGAAACGAGAGGTGGGGGAGACGAAGACCTGGAGAAATCAACATTCACCCGA 1620
 QY 1653 TCCCGAGAGGTGTGATGAGCAGAGGAGAGAGCTGTGCTGCTCCCGAGAGCT 1712
 Db 1621 TCCCGAGAGGTGTGATGAGCAGAGGAGAGAGCTGTGCTGCTCCCGAGAGCT 1680
 QY 1713 GAGAGACCAACGCTGTCTGAGAGAAAGGCGCTGAGAGACCCCAAGATGGAGAGCTGAG 1772
 Db 1681 GAGAGACCAACGCTGTCTGAGAGAAAGGCGCTGAGAGACCCCAAGATGGAGAGCTGAG 1740
 QY 1773 GAAGGAATCTCTCGATGAGAGAGAGAGAGAGATCTCCCTGGCATCTTC 1832
 Db 1741 GAAGGAATCTCTCGATGAGAGAGAGAGAGAGATCTCCCTGGCATCTTC 1800
 QY 1833 AAAAAAGATGTGACACCCAGAAACGGGTCCGAGACCTTCTGAGATGACAGAGAGGA 1892
 Db 1801 AAAAAAGATGTGACACCCAGAAACGGGTCCGAGACCTTCTGAGATGACAGAGAGGA 1860
 QY 1893 GAGCTGAGAGAGGTCAAGAGCCGACCTGTGCTCTCACTGATAGACAGTGTCAAGAAATG 1952
 Db 1861 GAGCTGAGAGAGGTCAAGAGCCGACCTGTGCTCTCACTGATAGACAGTGTCAAGAAATG 1920
 QY 1953 CAAGATGAATCAAAACTGTTGTGAGAGAACAAAAGCCAGAGAAACAAAGCTTAGGGTG 2012
 Db 1921 CAAGATGAATCAAAACTGTTGTGAGAGAACAAAAGCCAGAGAAACAAAGCTTAGGGTG 1980
 QY 2013 GATATCTAGTGTCTTGGAGACACTGATTTGTGTGCTGATCTCAAGAGAGAGACAAAG 2072
 Db 1981 GATATCTAGTGTCTTGGAGACACTGATTTGTGTGCTGATCTCAAGAGAGAGACAAAG 2040
 QY 2073 AAGGATCTCTCTTCAAGATGATGAGAGAGGCGCAAGAGACCTGGAGGGGACGTACAGA 2132
 Db 2041 AAGGATCTCTCTTCAAGATGATGAGAGAGGCGCAAGAGACCTGGAGGGGACGTACAGA 2100
 QY 2133 GCAGAGAGGCGCAGCAAAAGCAAAAGACCGGAGACAGACGCTGTTCTGCGACACCCAG 2192
 Db 2101 GCAGAGAGGCGCAGCAAAAGCAAAAGACCGGAGACAGACGCTGTTCTGCGACACCCAG 2160
 QY 2193 GAGCAGGACCAAGCGCAGAGAGTTCTTCAACCCGAGCCAGCGGAGAGCCCTTCCGAAGGG 2252
 Db 2161 GAGCAGGACCAAGCGCAGAGAGTTCTTCAACCCGAGCCAGCGGAGAGCCCTTCCGAAGGG 2220
 QY 2253 GAAGGTGTCTCTCACTTGGAGTCAATTTAAAGATTAGTCACTCCAAAGAAAAATCCAAAG 2312
 Db 2221 GAAGGTGTCTCTCACTTGGAGTCAATTTAAAGATTAGTCACTCCAAAGAAAAATCCAAAG 2280
 QY 2313 TCAAAATCTGAAGAGAAAGCCGAGAACTCTAGTGTAGACAGTTGTCCACTGAGATCGAA 2372
 Db 2281 TCAAAATCTGAAGAGAAAGCCGAGAACTCTAGTGTAGACAGTTGTCCACTGAGATCGAA 2340
 QY 2373 CCGAGTAGAGAAAGATCTTTGGGTTCTTAAAGAAATTCATCCCGAGAGGGGAGAGAA 2432
 Db 2341 CCGAGTAGAGAAAGATCTTTGGGTTCTTAAAGAAATTCATCCCGAGAGGGGAGAGAA 2400
 QY 2433 AGGAGAGACGGAGACAGAAACAGACACTGTGAGAAAGTCACTGAGGCGCACTGAGATTAAT 2492
 Db 2401 AGGAGAGACGGAGACAGAAACAGACACTGTGAGAAAGTCACTGAGGCGCACTGAGATTAAT 2460
 QY 2493 GAGAGACGACCTTAATGTCCAGACCGTGTGCTGTCTGAGATTAATGCACTGAGAGAG 2552
 Db 2461 GAGAGACGACCTTAATGTCCAGACCGTGTGCTGTCTGAGATTAATGCACTGAGAGAG 2520
 QY 2553 GAGAGAGATGAAAGCCAGGGGAAATCAGAGCTGCCAGACTGTGGGGCTGTGATCGTG 2612

Db 2521 GAGAGATGAAAGCCAGGGGAAATACGAGAGTGCCTCCAGCTGTGGGGCTGTGTACGTG 2580
 QY 2613 TCCGAGAGCTCAATTAAGACTGTGTCACACTGTGAGTGTCCAGATCATTTAGTGGAGC 2672
 Db 2581 TCCGAGAGCTCAATTAAGACTGTGTCACACTGTGAGTGTCCAGATCATTTAGTGGAGC 2640
 QY 2673 AGGCACTGTACCAATGTCTGAAGAGCGGTCTCTTGTGTGATATCCGCTTCCGTAAACGAA 2732
 Db 2641 AGGCACTGTACCAATGTCTGAAGAGCGGTCTCTTGTGTGATATCCGCTTCCGTAAACGAA 2700
 QY 2733 CCTCTTGAACACAGCGGGGAGAGACCATGCGACCTGTGAGAGGTCACTGAAAAAGAC 2792
 Db 2701 CCTCTTGAACACAGCGGGGAGAGACCATGCGACCTGTGAGAGGTCACTGAAAAAGAC 2760
 QY 2793 ATCAATTGAGAAAGAACTCTGTGTCTCAACCCAGACGTTACAGAGGGTAAAGATGCCAT 2852
 Db 2761 ATCAATTGAGAAAGAACTCTGTGTCTCAACCCAGACGTTACAGAGGGTAAAGATGCCAT 2820
 QY 2853 GACGACATGTCTACCACTGAGAGTGAATTTCACTCAGAGAGTGTGACACGACCAAGAGC 2912
 Db 2821 GACGACATGTCTACCACTGAGAGTGAATTTCACTCAGAGAGTGTGACACGACCAAGAGC 2880
 QY 2913 TCAGAGCTCTCCGTACTGAAGAGTTACCGAGCATCGGGGCGGAGAGACCAAGAC 2972
 Db 2881 TCAGAGCTCTCCGTACTGAAGAGTTACCGAGCATCGGGGCGGAGAGACCAAGAC 2940
 QY 2973 ATGTGTCCGAGATTTCCAGCTGACCTGCCAGACACCAAGAGAGAGCCACCA 3032
 Db 2941 ATGTGTCCGAGATTTCCAGCTGACCTGCCAGACACCAAGAGAGAGCCACCA 3000
 QY 3033 GTTCAGAGAGTGAAGAGTGTGTCTGTAGATTCAGAGAGAGAGAGCGCCAGAGCGAGCC 3092
 Db 3001 GTTCAGAGAGTGAAGAGTGTGTCTGTAGATTCAGAGAGAGAGAGCGCCAGAGCGAGCC 3060
 QY 3093 ATCTCCAGCGCTTGGACAGCAAGGTGAAGAGAGTCCGAGTGTCTGCAACCAAGCT 3152
 Db 3061 ATCTCCAGCGCTTGGACAGCAAGGTGAAGAGAGTCCGAGTGTCTGCAACCAAGCT 3120
 QY 3153 GTGCAGAGAACGGGGTCAAAAAGCACTGAGAGAGTTGAGAGAGTGAAGAGACTCCGAA 3212
 Db 3121 GTGCAGAGAACGGGGTCAAAAAGCACTGAGAGAGTTGAGAGAGTGAAGAGACTCCGAA 3180
 QY 3213 GTGCTGGCTTGGAGAGAGAGAGAGAGAGTATGCCGAAAGAGAGCCGTCAGAGAGCTGGA 3272
 Db 3181 GTGCTGGCTTGGAGAGAGAGAGAGAGAGTATGCCGAAAGAGAGCCGTCAGAGAGCTGGA 3240
 QY 3273 GCTGAGACTCTTGGACAGGGGCTGTGAGACTGAGACAGGCTACTCCAGAGAGCTTGAAGTT 3332
 Db 3241 GCTGAGACTCTTGGACAGGGGCTGTGAGACTGAGACAGGCTACTCCAGAGAGCTTGAAGTT 3300
 QY 3333 CCTGAAGTCAACGAGAGATGTAGACCATGTGCGCACGTGCCAGTTATCAAGCTCCAGAG 3392
 Db 3301 CCTGAAGTCAACGAGAGATGTAGACCATGTGCGCACGTGCCAGTTATCAAGCTCCAGAG 3360
 QY 3393 CTGATGGAACAGGCGGTGGCCCTGAGTCACTCCGAAACCTTGAACAGACAGTGAACAAT 3452
 Db 3361 CTGATGGAACAGGCGGTGGCCCTGAGTCACTCCGAAACCTTGAACAGACAGTGAACAAT 3420
 QY 3453 GGAAGCACTCCCTTAGAGAGATTGAGACACTGAGACCTGAGAGACAGCAAGAGAAACCAAT 3512
 Db 3421 GGAAGCACTCCCTTAGAGAGATTGAGACACTGAGACCTGAGAGACAGCAAGAGAAACCAAT 3480
 QY 3513 GACAGCCAGAGACAGTAAAGCCACTGACAGCTGTCAAGAGTCAAGAGTCAAGAGAGAG 3572
 Db 3481 GACAGCCAGAGACAGTAAAGCCACTGACAGCTGTCAAGAGTCAAGAGTCAAGAGAGAG 3540
 QY 3573 GCGGCTAATGCTCAGAGAGAGAGAGCTTCAACACTAATTAATGTTCCAGCCAGAGAA 3632
 Db 3541 GCGGCTAATGCTCAGAGAGAGAGAGCTTCAACACTAATTAATGTTCCAGCCAGAGAA 3600
 QY 3633 GAAACATGGGAGAGAACAGAGAGAGATGTTCTGAACCTACAGCAAGACCTTACTGCT 3692

Db 3601 GAACATGGGAGAAACAGGAGAGATGTTCTTGAACCTACACAGCAAGACTTACTGCT 3660
Qy 3693 GCAGCCGTGCCGTTCTGGCAAGACTGAGGTGGTCAAGAGGGTGAGTTGACTGGTTG 3752
Db 3661 GCAGCCGTGCCGTTCTGGCAAGACTGAGGTGGTCAAGAGGGTGAGTTGACTGGTTG 3720
Qy 3753 GATGGAGAAAAGTCAAGAGACAGAGGTGTTTGTACACTCTGACCCCAACAGTCAA 3812
Db 3721 GATGGAGAAAAGTCAAGAGAACAGAGGTGTTTGTACACTCTGACCCCAACAGTCAA 3780
Qy 3813 AAGCTCTCTGATGTGACATATGACAGTGAAGTGAATGGAGTGCCGGGTGTCAGGAAAAG 3872
Db 3781 AAGCTCTCTGATGTGACATATGACAGTGAAGTGAATGGAGTGCCGGGTGTCAGGAAAAG 3840
Qy 3873 GAGAGTACTGAAGTGCAGAGTCTTAGCCTCGAGAGGAGAGATGGAATCTGACGTTGAA 3932
Db 3841 GAGAGTACTGAAGTGCAGAGTCTTAGCCTCGAGAGGAGAGATGGAATCTGACGTTGAA 3900
Qy 3933 AAGGAGAAAGGAGACAAAGCCAGACAGTGAAGTGAAGAGGTGAGCAGGAACAGCC 3992
Db 3901 AAGGAGAAAGGAGACAAAGCCAGACAGTGAAGTGAAGAGGTGAGCAGGAACAGCC 3960
Qy 3993 GCTCCTGAGCATGAAGAACTTACGGAAGCCAGTCTGTACACTTGTACATGCCAGCTCA 4052
Db 3961 GCTCCTGAGCATGAAGAACTTACGGAAGCCAGTCTGTACACTTGTACATGCCAGCTCA 4020
Qy 4053 GAGAGGGGAGGACACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGACAA 4112
Db 4021 GAGAGGGGAGGACACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGACAA 4080
Qy 4113 GCAGTTGCATAGAGTTCAAGTTCAAAGCCTGAGACAAACAGTCACTCAAACAGCAGAA 4172
Db 4081 GCAGTTGCATAGAGTTCAAGTTCAAAGCCTGAGACAAACAGTCACTCAAACAGCAGAA 4140
Qy 4173 GCTGTGAAAAGTCTAGAAAAGTGTGATTTTCAGAGACAGTGAAGTCCAGAGTGT 4232
Db 4141 GCTGTGAAAAGTCTAGAAAAGTGTGATTTTCAGAGACAGTGAAGTCCAGAGTGT 4200
Qy 4233 GTAGGTGCACACTTATTACAGCTGAGAAGTCTCTGCAACGGGTGGCCACTGGACTCTT 4292
Db 4201 GTAGGTGCACACTTATTACAGCTGAGAAGTCTCTGCAACGGGTGGCCACTGGACTCTT 4260
Qy 4293 CAGCATGCAGAGACACGGTACCCCTGGGCGTGAAGTCTCAGGAGAAATCCATCCCAATC 4352
Db 4261 CAGCATGCAGAGACACGGTACCCCTGGGCGTGAAGTCTCAGGAGAAATCCATCCCAATC 4320
Qy 4353 ATAGTAACCTCTGCTCTGAAAACACCTACATCCTGACCTCAAGAGAGAAATAGCGCA 4412
Db 4321 ATAGTAACCTCTGCTCTGAAAACACCTACATCCTGACCTCAAGAGAGAAATAGCGCA 4380
Qy 4413 TCCAGAGAGAGCGATCAGAGGAGAGGACAAAGCCAGATGCTGCTCTGATGCTGACGGC 4472
Db 4381 TCCAGAGAGAGCGATCAGAGGAGAGGACAAAGCCAGATGCTGCTCTGATGCTGACGGC 4440
Qy 4473 AAGGAGGTACAGCAATCGAAAAGTCTCAAGGCTGAACCTGAGATCTTGGAACTTGAG 4532
Db 4441 AAGGAGGTACAGCAATCGAAAAGTCTCAAGGCTGAACCTGAGATCTTGGAACTTGAG 4500
Qy 4533 AGTAAGAGCAACAGATTTGCTGMAAGTCAATTCAGACAGCCGCTGACCAAGTTCCGACGT 4592
Db 4501 AGTAAGAGCAACAGATTTGCTGMAAGTCAATTCAGACAGCCGCTGACCAAGTTCCGACGT 4560
Qy 4593 ACAGAAACAGCCCCCGAACTCATGTTATGATTTACAGACCCAGGTTCTCTGATGCGAGG 4652
Db 4561 ACAGAAACAGCCCCCGAACTCATGTTATGATTTACAGACCCAGGTTCTCTGATGCGAGG 4620
Qy 4653 CTTGACAGCAGGAGGCCAAACAGATGCTGGACAAAATGAAGATGCAAGATGAACAC 4712
Db 4621 CTTGACAGCAGGAGGCCAAACAGATGCTGGACAAAATGAAGATGCAAGATGAACAC 4680
Qy 4713 CCAGTCCGAGCCGACAGAGGACTTGAAGTCTGACCGCTTCTGAGGAGATGGGCTCAG 4772
Db 4681 CCAGTCCGAGCCGACAGAGGACTTGAAGTCTGACCGCTTCTGAGGAGATGGGCTCAG 4740

RESULT 7

PCT-US97-06830-1

; Sequence 1, Application PC/TUS9706830

; GENERAL INFORMATION:

; APPLICANT: Gelman, Irwin H.

; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10112-0228

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US97/06830

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/665,401

; FILING DATE: 18-JUN-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Richard S

; REGISTRATION NUMBER: 26,154

; REFERENCE/DOCKET NUMBER: A30558 - 165/34008

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-408-2558

; TELEFAX: 212-765-2519

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5134 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

Db 1968 TTTCCAGCTGACTGCTCCCGACACACACAGAGGAGCCACCCAGTTTCAGAGGTAG 2027
Qy 3046 AGAGTGTGTCTAGATACAGAGAGAGAGAGCGCCAGAGCGCCAGTCTCTCAAGCCG 3105
Db 2028 AGGTGTGTGTCTAGATACAGAGAGAGAGAGCGCCAGAGCGCCAGTCTCTCAAGCCG 2087
Qy 3106 TTGCAGACAAAGGTGAAGAGAGGTCCAGGTGCTGCTCAACCCAGACGTGTCAGAGAACGG 3165
Db 2088 TTGCAGACAAAGGTGAAGAGAGGTCCAGGTGCTGCTCAACCCAGACGTGTCAGAGAACGG 2147
Qy 3166 GGTCAAAAGCACTGGAGAGGTGAGAGAGGTAGAGAGGAGTCCGAAAGTGTGCTTCGG 3225
Db 2148 GGTCAAAAGCACTGGAGAGGTGAGAGAGGTAGAGAGGAGTCCGAAAGTGTGCTTCGG 2207
Qy 3226 AGAAAGAGAGAGCTTATGCCGAAAGGACCCGTGACAGAGCTGAGAGCTGAGCATTTG 3285
Db 2208 AGAAAGAGAGAGCTTATGCCGAAAGGACCCGTGACAGAGCTGAGAGCTGAGCATTTG 2267
Qy 3286 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCTTGAAGTTCCCTGAGTCAAG 3345
Db 2268 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCTTGAAGTTCCCTGAGTCAAG 2327
Qy 3346 CAGATGTAGACCTATGCTCCGACAGTGTATCAAGCTCCAGCAGCTGATGGAACAGG 3405
Db 2328 CAGATGTAGACCTATGCTCCGACAGTGTATCAAGCTCCAGCAGCTGATGGAACAGG 2387
Qy 3406 CCGTGGCCCTGAGTATCCGAAACCTTTGACAGACAGTGAACAAATGGAAGCACTCCCT 3465
Db 2388 CCGTGGCCCTGAGTATCCGAAACCTTTGACAGACAGTGAACAAATGGAAGCACTCCCT 2447
Qy 3466 TAGCAGATTCAGACACTGCAGATGGGACACAGCAGATGAACCCATTGACAGCCAGACA 3525
Db 2448 TAGCAGATTCAGACACTGCAGATGGGACACAGCAGATGAACCCATTGACAGCCAGACA 2507
Qy 3526 GTAAAGCACTGCAGCTGTGAGGAGTCAAGGTCACAGAGAGAGCGCTACTGCTC 3585
Db 2508 GTAAAGCACTGCAGCTGTGAGGAGTCAAGGTCACAGAGAGAGCGCTACTGCTC 2567
Qy 3586 AGAAAGAGAGCTTCGACACTACTTAATATGTTCCAGCCCGAGGAACATGGGGAAG 3645
Db 2568 AGAAAGAGAGCTTCGACACTACTTAATATGTTCCAGCCCGAGGAACATGGGGAAG 2627
Qy 3646 AACAGAGAGATGTTCTTGAACCTACAGCAGAGAGCTTACTGCTGACCGCTGCCCG 3705
Db 2628 AACAGAGAGATGTTCTTGAACCTACAGCAGAGAGCTTACTGCTGACCGCTGCCCG 2687
Qy 3706 TTCTGGCAAAAGACTGAGTGGGTCAAGAGGTGAGTTGACTGTTGATGGAGAAAG 3765
Db 2688 TTCTGGCAAAAGACTGAGTGGGTCAAGAGGTGAGTTGACTGTTGATGGAGAAAG 2747
Qy 3766 TCAAGAGAACAGAGGTGTTTGTACACTCTGGACCCAAAGCTCAAAAGGCTGCTGATG 3825
Db 2748 TCAAGAGAACAGAGGTGTTTGTACACTCTGGACCCAAAGCTCAAAAGGCTGCTGATG 2807
Qy 3826 TGACATATGACAGTGAAGTATGGAGTGGCGGGTGTAGGAAAGAGAGAGTACTGAAG 3885
Db 2808 TGACATATGACAGTGAAGTATGGAGTGGCGGGTGTAGGAAAGAGAGAGTACTGAAG 2867
Qy 3886 TGACAGTCTTAGCTCGAGAGGAGAGATGGAACCTGACGTTGAAAGAGAGAAAGGG 3945
Db 2868 TGACAGTCTTAGCTCGAGAGGAGAGATGGAACCTGACGTTGAAAGAGAGAAAGGG 2927
Qy 3946 AGACAAAGCCAGAGCAAGTGAAGAGGTGAGCAGGAGAAACAGCCGCTCCTGAGCATG 4005
Db 2928 AGACAAAGCCAGAGCAAGTGAAGAGGTGAGCAGGAGAAACAGCCGCTCCTGAGCATG 2987
Qy 4006 AAGGAACCTACGGAAGCCAGTCTGACATTTGATGCCAGCTCAGAGAGGGGAGG 4065
Db 2988 AAGGAACCTACGGAAGCCAGTCTGACATTTGATGCCAGCTCAGAGAGGGGAGG 3047
Qy 4066 CACTGGAGAGCTTGGAGAGCCCTTCTCTCCAGACCAAGACAAAGAGGTGTCATG 4125
Db 3048 CACTGGAGAGCTTGGAGAGCCCTTCTCTCCAGACCAAGACAAAGAGGTGTCATG 3107

Qy 4126 AGTTTCAAGTTTCAAGCCCTGGACACAAAGTCACTCAAAACAGCAGAAAGCTGTGAAAAGG 4185
Db 3108 AGTTTCAAGTTTCAAGCCCTGGACACAAAGTCACTCAAAACAGCAGAAAGCTGTGAAAAGG 3167
Qy 4186 TCATAGAAACGGTTGTGATTTTCAGAGACAGGTGAAAAGTGTAGAGTGTAGAGTGCACACT 4245
Db 3168 TCATAGAAACGGTTGTGATTTTCAGAGACAGGTGAAAAGTGTAGAGTGTAGAGTGCACACT 3227
Qy 4246 TATTACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGG 4305
Db 3228 TATTACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGG 3287
Qy 4306 ACAGGTATCCCTTGGGCTGAGTCTCAGCAGAAATCCATCCCAATCATAGTAATCTCTG 4365
Db 3288 ACAGGTATCCCTTGGGCTGAGTCTCAGCAGAAATCCATCCCAATCATAGTAATCTCTG 3347
Qy 4366 CTCCTCAAAAGCACCTTACATCTCTGACCTCAAGGAGAAATAGCGCATCCAGAGAGAGC 4425
Db 3348 CTCCTCAAAAGCACCTTACATCTCTGACCTCAAGGAGAAATAGCGCATCCAGAGAGAGC 3407
Qy 4426 GATCAGAGGAGAGGACCAAGCCAGATGCTGCTGATGCTGACGGCAAGGAGAGTACAG 4485
Db 3408 GATCAGAGGAGAGGACCAAGCCAGATGCTGCTGATGCTGACGGCAAGGAGAGTACAG 3467
Qy 4486 CAATCGAAAAGTCTCAAGGCTGAACTCTGAGATCTCTGAACTTTGAGAGTGAAGCAACA 4545
Db 3468 CAATCGAAAAGTCTCAAGGCTGAACTCTGAGATCTCTGAACTTTGAGAGTGAAGCAACA 3527
Qy 4546 AGATTGCTGTGAACGTCATTTCAGACAGCGTTCAGCAGTTCGCACGTACAGAACAGCCC 4605
Db 3528 AGATTGCTGTGAACGTCATTTCAGACAGCGTTCAGCAGTTCGCACGTACAGAACAGCCC 3587
Qy 4606 CCGAAACTCATGCTTATGATTCACAGACCCAGGTTCTGCTGATGACAGCTTCAGACAGGG 4665
Db 3588 CCGAAACTCATGCTTATGATTCACAGACCCAGGTTCTGCTGATGCTGCAATGCGCTTGGACAGAGG 3647
Qy 4666 AGCCCAACAGATGCTGGACAAAATGAAAGTGCAGAGTGCAGAAACCCAGTCCCGCAGC 4725
Db 3648 AGCCCAACAGATGCTGGACAAAATGAAAGTGCAGAGTGCAGAAACCCAGTCCCGCAGC 3707
Qy 4726 CCAGAGAGACTTGCAGTTCCTGACCGTTCCTGAGGAGCATGGGCTCAGCTCGAATGCT 4785
Db 3708 CCAGAGAGACTTGCAGTTCCTGACCGTTCCTGAGGAGCATGGGCTCAG - CTCGAAATGCT 3765
Qy 4786 TGCCGCGCTTGCAAGTTGAAAAGCCCGTGTCAAAGTGAAGTGAAGAGCTGCTCTCA 4845
Db 3766 TGCCGCGCTTGCAAGTTGAAAAGCCCGTGTCAAAGTGAAGTGAAGAGCTGCTCTCA 3825
Qy 4846 ACCCAAGATCCAAAAGGAGCATGCTGCTGATGGCCCTCAGTCCAAAAGCTTAGCCAGG 4905
Db 3826 ACCCAAGAT - CAAAAGGAGCATGCTGCTGATGGCCCTCAGTCCAAAAGCTTAGCCAGG 3884
Qy 4906 CAGAGCCAGTGCCTCTGAAAACCTTAACCAAGAAATCCCCAGACACCAACGGACCAAGC 4965
Db 3885 CAGAGCA - - - GTGTCTGAAAACCTTAACCAAGAAATCCCCAGACACCAACGGACCAAGC 3941
Qy 4966 TAACCCAGAGGCGCATGCCCAAAAGTTGAGTCCAGAGAAAGAAATGCTTACCAAGT 5025
Db 3942 TAACCCAGAGGCG - ATGCCCCCAAAAGTTGAGTCCAGAGAAAGAAATGCTTACCAAGT 4000
Qy 5026 CAGTCAAAAGAGAACAAAGGCCAGGACAGAGGAGCTTCAGAGGAGGAGGAGGAGCTGG 5085
Db 4001 CAGTCAAAAGAGAACAAAGGCCAGGACAGAGGAGCTTCAGAGGAGGAGGAGGAGCTGG 4060
Qy 5086 CAGAAATCTTAAGATGTTAGTGTCTTATGATCTGTAAGACCAAGATGTGAAAACAAG 5145
Db 4061 CAGAAATCTTAAGATGTTAGTGTCTTATGATCTGTAAGACCAAGATGTGAAAACAAG 4120
Qy 5146 TCACAGAACAAAGATGCTGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCATGATC 5205
Db 4121 TCACAGAACAAAGATGCTGCTGTTGGGACCTTG - GACCAAGATTTTCAGAGCCATGATC 4179

```

QY 5206 CAGAGAGAGGCGCTCCATGATTTTCCACCGAGAGCAACCCGCAATTTTGAGGCT 5265
DB 4180 CAGAGAGAGGCGCTCCATGATTTTCCACCGAGAGCAACCCGCAATTTTGAGGCT 4239
QY 5266 TCATCGGAGCTAGAGCCAGCTAAATTTCTGTTTCAAGCTGCTTTGATTTGGCC 5325
DB 4240 TCATCGGAGCTAGAGCCAGCTAAATTTCTGTTTCAAGCTGCTTTGATTTGGCC 4299
QY 5326 TTGATGCCGCTCGTGTATTTCTPAACATTTCTGTTTCAAGCTGCTTTGATTTGGCC 5385
DB 4300 TTGATGCCGCTCGTGTATTTCTPAACATTTCTGTTTCAAGCTGCTTTGATTTGGCC 4359
QY 5386 TTGATGCCGCTCGTGTATTTCTGATTTAAAGTCTGCTTCAACCTGGAACCAATTC 5445
DB 4360 TTGATGCCGCTCGTGTATTTCTGATTTAAAGTCTGCTTCAACCTGGAACCAATTC 4419
QY 5446 GCCATACCTAGTTCACTTCTCAACTGGAGATCTCTTATGATTTATGATGAT 5505
DB 4420 GCCATACCTAGTTCACTTCTCAACTGGAGATCTCTTATGATTTATGATGAT 4479
QY 5506 TTATGATGCTCTCTCTCTGATCTATTTGATTTTCTTCAAGCTTTAAGCAATGCT 5565
DB 4480 TTATGATGCTCTCTCTCTGATCTATTTGATTTTCTTCAAGCTTTAAGCAATGCT 4539
QY 5566 TTTTGTATTGCAATATATTAACGGGTGTGACGCAATGCGGCTTTGAAAAGCTCCAA 5625
DB 4540 TTTTGTATTGCAATATATTAACGGGTGTGACGCAATGCGGCTTTGAAAAGCTCCAA 4599
QY 5626 GGCTCAACGTGTAACCTGAGCAAGATACATTCCTGGAAGAGAGCAAGTCTTTT 5685
DB 4600 GGCTCAACGTGTAACCTGAGCAAGATACATTCCTGGAAGAGAGCAAGTCTTTT 4659
QY 5686 TTAAGTTACTGATGCTTATGATCTGTGAGCTTCTAGTCTCTGAAAAGTGTGTTTCC 5745
DB 4660 TTAAGTTACTGATGCTTATGATCTGTGAGCTTCTAGTCTCTGAAAAGTGTGTTTCC 4719
QY 5746 TATGACACGAGCTCAGAAATTAACCCCATTTTGAACATCCAGATGTCCCAATTC 5805
DB 4720 TATGACACGAGCTCAGAAATTAACCCCATTTTGAACATCCAGATGTCCCAATTC 4779
QY 5806 TACCATGATTTTCCCTCTTTTGTCTATGCAATGCAAGTGGAAAGAGCTCTCT 5865
DB 4780 TACCATGATTTTCCCTCTTTTGTCTATGCAATGCAAGTGGAAAGAGCTCTCT 4839
QY 5866 GTGTGATTAAGCCCTGTCTCTTAATGATATGACAAATGAGTGTGCTTAAGCCCATGA 5925
DB 4840 GTGTGATTAAGCCCTGTCTCTTAATGATATGACAAATGAGTGTGCTTAAGCCCATGA 4899
QY 5926 GATGTTTCTTAATGCAAGAGATCTGTGTGATGCTTTTGTATTTACTCTTTCTATGCT 5985
DB 4900 GATGTTTCTTAATGCAAGAGATCTGTGTGATGCTTTTGTATTTACTCTTTCTATGCT 4959
QY 5986 GGACCGAATTCATATGCAAGATGAGTGAATGCTGTTCTTTTCAAGATGATTTGATAG 6045
DB 4960 GGACCGAATTCATATGCAAGATGAGTGAATGCTGTTCTTTTCAAGATGATTTGATAG 5019
QY 6046 ATACTGAGTTTGTCTGTATATCTGTGCTGCTTTTCAAGAACATGTGCAATATG 6105
DB 5020 ATACTGAGTTTGTCTGTATATCTGTGCTGCTTTTCAAGAACATGTGCAATATG 5079
QY 6106 TTCTTTGATTAATTTGTGATTTGACACTGATTTAAATTAACATATTTGACTAC 6160
DB 5080 TTCTTTGATTAATTTGTGATTTGACACTGATTTAAATTAACATATTTGACTAC 5134

```

RESULT 8
US-08-665-401-1

Sequence 1, Application US/08665401
GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

```

ADDRESS: Brunbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,401
FILING DATE: 18-JUN-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5134 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-665-401-1

```

Query Match 78.4%; Score 4831.4; DB 10; Length 5134;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 5048; Conservative 0; Mismatches 76; Indels 31; Gaps 14;

```

QY 1016 GGAAGAAGCAAGCCAGCTCTGAGAGAGCAGAGCCGCGCAAGACACAGCCAGGCTAG 1075
DB 1 GGAAGAAGCAAGCCAGCTCTGAGAGAGCAGAGCCGCGCAAGACACAGCCAGGCTAG 60
QY 1076 GTTGTACAGCACTACAGAGAGTGTGCTCTTTTGAAGACCAAGTGTGAGACTTGA 1135
DB 61 GTTGTACAGCACTACAGAGAGTGTGAGTGTGCTCTTTTGAAGACCAAGTGTGAGACTTGA 120
QY 1136 GGCATGTCAGAGAGAGAGTGTGCTCTTTTGAAGACCAAGTGTGAGAGATGGA 1195
DB 121 GGCATGTCAGAGAGAGAGTGTGCTCTTTTGAAGACCAAGTGTGAGAGATGGA 180
QY 1196 AGCCACCAAGAGTGTGTCAGAGAGTCAAGTGAACACCTGTGAGAGAGAGAGAGAG 1255
DB 181 AGCCACCAAGAGTGTGTCAGAGAGTCAAGTGAACACCTGTGAGAGAGAGAGAGAG 240
QY 1256 GCAGGAGAGAGAGAGAGGCTGAAGGGGGGTGTGATGAGAGAGAGAGAGAGATCTT 1315
DB 241 GCAGGAGAGAGAGAGAGGCTGAAGGGGGGTGTGATGAGAGAGAGAGAGATCTT 300
QY 1316 GCCCCTGAGAACTGCTGAGCCCGCAGAGAGTCCCGCAGAGAGCTGAGTGAAGA 1375
DB 301 GCCCCTGAGAACTGCTGAGCCCGCAGAGAGTCCCGCAGAGAGCTGAGTGAAGA 360
QY 1376 GCTGATGAAGACAGAGAGATGTGTCTCTGAGAGAGACCACTCACTGACAGACT 1435
DB 361 GCTGATGAAGACAGAGAGATGTGTCTCTGAGAGAGACCACTCACTGACAGACT 420
QY 1436 AAGTCCTGAAGAGAGAGCTGCCCAACCCAGAGAGCAATTCAGTGAAGTGAAGAT 1495

```


Db 421 AAGTCTGAGAGAGAGCGCTGCCAAACACCCAGAGGCAATTGTCAGTGAGGTGGAGAT 480
Qy 1496 GCTGTCTCTCAGAGAAAGATCAAGGTACAGGGAAAGTCCCTTGAAGAAATCTTTCACTAG 1555
Db 481 GCTGTCTCTCAGAGAAAGATCAAGGTACAGGGAAAGTCCCTTGAAGAAATCTTTCACTAG 540
Qy 1556 CTGAGGCTTAAAGAAAGCTGTCTGGAAAGACAGAGGGGAAACAGAGAGGTGGGGAGA 1615
Db 541 CTGAGGCTTAAAGAAAGCTGTCTGGAAAGACAGAGGGGAAACAGAGAGGTGGGGAGA 600
Qy 1616 CGAAGAGCTGGAGAAATACCAACACATTTACACCCGAATCCCCAGAGAGTGCTGATGAGCA 1675
Db 601 CGAAGAGCTGGAGAAATACCAACACATTTACACCCGAATCCCCAGAGAGTGCTGATGAGCA 660
Qy 1676 GAAGGGAGAGAGCTCTCGTGTCTCCCGAGAGAGCTTGAGAGACCAAGTGTCTGGAGAA 1735
Db 661 GAAGGGAGAGAGCTCTCGTGTCTCCCGAGAGAGCTTGAGAGACCAAGTGTCTGGAGAA 720
Qy 1736 AGGCGCTGGAGACACCCAGAGATGGGAGCTGAGAGAGGAACTACTTCCGATGGAGA 1795
Db 721 AGGCGCTGGAGACACCCAGAGATGGGAGCTGAGAGAGGAACTACTTCCGATGGAGA 775
Qy 1796 GAAGAGAGAGAGAGGATCACTCCCTGGGATCCCTTCAAAAAGATGGTGACACCCAGAA 1855
Db 776 GAGAGAGAGAGAGGATCACTCCCTGGGATCCCTTCAAAAAGATGGTGACACCCAGAA 835
Qy 1856 ACGGCTCCGAAGACCTTTCTGAGAGTGACAAAGGAGAGAGCTGGAGAGGTCAAGAGCGC 1915
Db 836 AC-GGTCGGAAGACCTTTCTGAGAGTGACAAAGGAGAGAGCTGGAGAGGTCAAGAGCGC 894
Qy 1916 CAGTTGTCTCCACTGATAGCACAGTGTGAGAAATCCAGATGAAGTCAAACTGTTGG 1975
Db 895 CAGTTGTCTCCACTGATAGCACAGTGTGAGAAATCCAGATGAAGTCAAACTGTTGG 954
Qy 1976 TGAGGAACAAAGCCAGAGAAACCAAGCGTGTAGGTCGATCTTCACTGTCTTGGGAAGC 2035
Db 955 TGAGGAACAAAGCCAGAGAAACCAAGCGTGTAGGTCGATCTTCACTGTCTTGGGAAGC 1014
Qy 2036 ACTGATTTGTGTCGATCATCCAGAGAGAGCAAGAAAGGATCTCTTCAAGATGATGA 2095
Db 1015 ACTGATTTGTGTCGATCATCCAGAGAGAGCAAGAAAGGATCTCTTCAAGATGATGA 1071
Qy 2096 AGAGGGCCAGGACACTGGGA-GGGACAGTACAGAGCAGAGAGGAGCCAGCAAGACA 2154
Db 1072 AAGAGGCCAAAGGACACTGGGAGGGGACAGTCAAGAGAGAGGAGCCAGCAAGACA 1131
Qy 2155 AAGAAGCCGAAACAGACGCTGTTCTCCAGACCCAGAGAGAGGACCAAGCGCAAGGA 2214
Db 1132 AAGAGCC-GAAGAGAGTGTCTCTCCAGACCCAGAGAGAGGACCAAGCGCAAGGA 1190
Qy 2215 GTTCTCACCAGCCAGCGGAGAGCCCTTCCGAAGGGGAGGTGTCTCCACTTGGGAGT 2274
Db 1191 GTTCTCACCAGCCAGCGGAGAGCCCTTCCGAAGGGGAGGTGTCTCCACTTGGGAGT 1250
Qy 2275 CATTTAAAGATTAGTCACTCCAGAAAGAAATCCAAAGTCAAACTGGAAGAGAAAGCCG 2334
Db 1251 CATTTAAAGATTAGTCACTCCAGAAAGAAATCCAAAGTCAAACTGGAAGAGAAAG 1310
Qy 2335 AAGACTCTAGTGTAG-----AGCAGTGTCTCACTGAGATCGAACCCGATGAGAGAG 2386
Db 1311 CCGAAGGACTCTAGTGTAGGAGAGGTGTCTCACTGAGATCGAACCCGATGAGAGAG 1370
Qy 2387 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGGAGCGGAGAAAGAGGCGAGACGGAA 2446
Db 1371 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGGAGCGGAGAAAGAGGCGAGATGGAA 1430
Qy 2447 -GCAAGAACAGGCACTGTGGAGAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCTTA 2505
Db 1431 GGAAGAACAGGCACTGTGGAGAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCTTA 1490
Qy 2506 ATGTCCAGCGCTGTGCTCTCTGATGATATGTCAGTGGAGAGGAGAGATGGAAG 2565
Db 1491 ATGTCCAGCGCTGTGCTCTCTGATGATATGTCAGTGGAGAGGAGAGATGGAAG 1550

Qy 2566 CCAGGGGAATACGAGAGCTGCCCAAGCTGCTGGGGCTGTGTACGTGTCCGAGAGCTCA 2625
Db 1551 CCAGGGGAATACGAGAGCTGCCCAAGCTGCTGGGGCTGTGT---AGTGTCCGAGAGCTCA 1607
Qy 2626 GTAAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGACACAGGGCAGTCA 2685
Db 1608 GTAAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGACACAGGGCAGTCA 1667
Qy 2686 GTGTGAAAGAGCGGTCTCTCTTCGTGGATATCCCGTTCCGTAACAGAACCTTTTGAACACA 2745
Db 1668 GTGTGAAAGAGCGGTCTCTCTTCGTGGATATCCCGTTCCGTAACAGAACCTTTTGAACACA 1727
Qy 2746 CAGCGGAGAGGCAATGCCACCTGTTTGAAGAGGTCACTGAAAAGACATCAITTCGAGAAG 2805
Db 1728 CAGCGGAGAGGCAATGCCACCTGTTTGAAGAGGTCACTGAAAAGACATCAITTCGAGAAG 1787
Qy 2806 AAACCTCTGTGCTCACCCAGAGCTTACCAGAGGTTAAAGTCCCATGACGACATGGTCA 2865
Db 1788 AAACCTCTGTGCTCACCCAGAGCTTACCAGAGGTTAAAGTCCCATGACGACATGGTCA 1847
Qy 2866 CCAAGTGAAGTGAATTTACCTCAGAAAGCTGTGACAGCTCAGAGAGCTCAGAGGCTCTCC 2925
Db 1848 CCAAGTGAAGTGAATTTACCTCAGAAAGCTGTGACAGCTCAGAGAGCTCAGAGGCTCTCC 1907
Qy 2926 GTACTGAAGAGTGTACCGAAGCATCCGGGGCCGAAGAGACCAAGAGACATGGTGTCCGCGAG 2985
Db 1908 GTACTGAAGAGTGTACCGAAGCATCCGGGGCCGAAGAGACCAAGAGACATGGTGTCCGCGAG 1967
Qy 2986 TTTCCAGCTGACTGACTCTCCAGACACACAGAGAGGAGCCACCCAGTTTCAGAGGCTAG 3045
Db 1968 TTTCCAGCTGACTGACTCTCCAGACACACAGAGAGGAGCCACCCAGTTTCAGAGGCTAG 2027
Qy 3046 AGAGTGTGTGTGTAGATACAGAAAGAGAGGAGCCAGAGGAGGAGCCATCTCCAGAGCG 3105
Db 2028 AGGTTGTGTGTGTAGATACAGAAAGAGAGGAGCCAGAGGAGGAGCCATCTCCAGAGCG 2087
Qy 3106 TTGCAGACAAGGTGAAAGAGGAGTCCAGGTGCTCTCAACCCAGACTGTGCAGAGAACGG 3165
Db 2088 TTGCAGACAAGGTGAAAGAGGAGTCCAGGTGCTCTCAACCCAGACTGTGCAGAGAACGG 2147
Qy 3166 GGTCAAAAGCACTGGAGAAAGTTTGAAGAGGTAGAGAGAGTCCGAAAGTGTGCTGCTCG 3225
Db 2148 GGTCAAAAGCACTGGAGAAAGTTTGAAGAGGTAGAGAGGAGTCCGAAAGTGTGCTGCTCG 2207
Qy 3226 AGAAAGAGAGAGAGCTTATGCGGAAAGAGCCGTCGAGAGAGTGGAGCTGAGCATCTTG 3285
Db 2208 AGAAAGAGAGAGAGCTTATGCGGAAAGAGCCGTCGAGAGAGTGGAGCTGAGCATCTTG 2267
Qy 3286 CACAGGCTCTGAGACTGGAGAGGCTTCTCCAGAGAGGCTTGAAGTTCCTGAAGTCAAGC 3345
Db 2268 CACAGGCTCTGAGACTGGAGAGGCTTCTCCAGAGAGGCTTGAAGTTCCTGAAGTCAAGC 3327
Qy 3346 CAGATGTAGACATGTGCGCACGTGCGAGGTTATCAAGTCCAGAGAGTGTGAGAACAG 3405
Db 2328 CAGATGTAGACATGTGCGCACGTGCGAGGTTATCAAGTCCAGAGAGTGTGAGAACAG 2387
Qy 3406 CCGTGGCCCTGAGTCACTCCGAAACCTTGACAGACAGTGGAGCAATGGAAGCACTCCCT 3465
Db 2388 CCGTGGCCCTGAGTCACTCCGAAACCTTGACAGACAGTGGAGCAATGGAAGCACTCCCT 2447
Qy 3466 TAGCAGATTTCAGACACTGAGATGGGACACAGACAGATGAAACCTTTCAGAGCAGGACA 3525
Db 2448 TAGCAGATTTCAGACACTGAGATGGGACACAGACAGATGAAACCTTTCAGAGCAGGACA 2507
Qy 3526 GTAAAGCCACTGAGCTGTGAGGAGTCAAGGTCAAGAGAGAGGCGGCTACTGCTC 3585
Db 2508 GTAAAGCCACTGAGCTGTGAGGAGTCAAGGTCAAGAGAGAGGCGGCTACTGCTC 2567
Qy 3586 AGAAAGAGAGGCTTTCGACACTACCTTAATGTTTCCAGCCAGGAGAGCAATGGGGAAG 3645
Db 2568 AGAAAGAGAGGCTTTCGACACTACCTTAATGTTTCCAGCCAGGAGAGCAATGGGGAAG 2627


```
Db 4780 TACCATGATTTTTCCTCCCTTTTTCCTAATCCAGTCCAGGTTGGAAGAAGTCTCTCT 4839
Qy 5866 GTGTGAGATTAAGCCCTGTCTCTTAATGATATGACAAATGAGTGTGCTTAAGCCCATGA 5925
Db 4840 GTGTGAGATTAAGCCCTGTCTCTTAATGATATGACAAATGAGTGTGCTTAAGCCCATGA 4899
Qy 5926 GATGTTTCTTAATCCAGAGGAATCTGTTGACGTTTTTTTGGATTGTTACTTCTATGCT 5985
Db 4900 GATGTTTCTTAATCCAGAGGAATCTGTTGACGTTTTTTTGGATTGTTACTTCTATGCT 4959
Qy 5986 GGACCCGAATTCATATGACAGATCGAAGTGAGTCCCTGTTCTTTACAGATGGTATTTGATAG 6045
Db 4960 GGACCCGAATTCATATGACAGATCGAAGTGAGTCCCTGTTCTTTACAGATGGTATTTGATAG 5019
Qy 6046 ATACTGGAGTTTGTCTGTGTTATATCTGTGCCCTTCTTTAAGAACAAATGTTGCAATTATG 6105
Db 5020 ATACTGGAGTTTGTCTGTGTTATATCTGTGCCCTTCTTTAAGAACAAATGTTGCAATTATG 5079
Qy 6106 TTCCTTTGGATAAATTTGATTTGCACAACTGATTTAAATAAACAATATTTGACTAC 6160
Db 5080 TTCCTTTGGATAAATTTGATTTGCACAACTGATTTAAATAAACAATATTTGACTAC 5134

RESULT 9
US-09-902-432-1
; Sequence 1, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A, 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902.432
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5134
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-1

Query Match 78.4%; Score 4831.4; DB 34; Length 5134;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 5048; Conservative 0; Mismatches 76; Indels 31; Gaps 14;

Qy 1016 GGAAGAAGACAGAGCCAGCTCGGAGGAGCAGGAGCCGGCAGAAGACACAGACCAGGCCAG 1075
Db 1 GGAAGAAGACAGAGCCAGCTCGGAGGAGCAGGAGCCGGCAGAAGACACAGACCAGGCCAG 60

Qy 1076 GTTGTGACGACATACGAGAAAGTGGAGTCCCTTTTGGAAACACGAGTTGGTGCACCTGGA 1135
Db 61 GTTGTGACGACATACGAGAAAGTGGAGTCCCTTTTGGAAACACGAGTTGGTGCACCTGGA 120

Qy 1136 GGCATCGTCAGAGAGAGTGTCTCTTTGGCAACCGAAGTGTGATGAGAAGATGGA 1195
Db 121 GGCATCGTCAGAGAGAGTGTCTCTTTGGCAACCGAAGTGTGATGAGAAGATGGA 180

Qy 1196 AGCCCAACCAAGATTTGTCAGAGGTCACGTCGAGACCCGTGGAGAGACAGAGGAGGA 1255
Db 181 AGCCCAACCAAGATTTGTCAGAGGTCACGTCGAGACCCGTGGAGAGACAGAGGAGGA 240

Qy 1256 GCAGGGAGGAGGAGAGGCTGAAGGGGGCGTGTGTGTAGAGGAACAGAGAGATCCTT 1315
Db 241 GCAGGGAGGAGGAGAGGCTGAAGGGGGCGTGTGTGTAGAGGAACAGAGAGATCCTT 300

Qy 1316 GCCCCTTGAGAACTGGCTGAGCCCGCAGGAGGTCCTCCCGAGGAAGCTGAGCCTGCTGAGGA 1375
```

```
Db 301 GCCCCTTGAGAAACTGGCTGAGCCCCCAGGAGGTCCTCCAGGAAGCTGAGCCTGCTGAGGA 360
Qy 1376 GCTGATGAAGACAGCAGAGAGATGTGTCTCTGGAGGAGACCACATCAACTGACAGACCT 1435
Db 361 GCTGATGAAGACAGCAGAGAGATGTGTCTCTGGAGGAGACCACATCAACTGACAGACCT 420
Qy 1436 AAGTCTCTGAAGAGAACGCTGCCAAAACACCCAGAAAGGCAATTTGTCTAGTGAAGTGGAGAT 1495
Db 421 AAGTCTCTGAAGAGAACGCTGCCAAAACACCCAGAAAGGCAATTTGTCTAGTGAAGTGGAGAT 480
Qy 1496 GCTGTCTCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTGAAGAAAATCTTTCACTAG 1555
Db 481 GCTGTCTCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTGAAGAAAATCTTTCACTAG 540
Qy 1556 CTCAAGGCTTAAAGAAGCTGTCTGGGAAGAACAGAGAGGGGAAAACGAGGAGGTGGGGGAGA 1615
Db 541 CTCAAGGCTTAAAGAAGCTGTCTGGGAAGAACAGAGAGGGGAAAACGAGGAGGTGGGGGAGA 600
Qy 1616 CGAAGAGCCTGGAGAAATACCAACACATTTACACCGAATCTCCCAAGAGAGTCTGATGAGCA 1675
Db 601 CGAAGAGCCTGGAGAAATACCAACACATTTACACCGAATCTCCCAAGAGAGTCTGATGAGCA 660
Qy 1676 GAAGGGAGAGAGCTCTGCGTCTGCTCCCGGAGGAGCCTGAGGAGACCACGTCGTCTGGAGAA 1735
Db 661 GAAGGGAGAGAGCTCTGCGTCTGCTCCCGGAGGAGCCTGAGGAGACCACGTCGTCTGGAGAA 720
Qy 1736 AGGGCCGCTGGAAGCACTCCAGGATGGGGAAGCTGAGGAAGGAATCTCTCCGATGAGGA 1795
Db 721 AGGGCCGCTGGAAGCA-CCAGAGTGGGAAGCTGAGGAAGGAATCTCTCTCTGGA 775
Qy 1796 GAAGAAGAGAGAGGAGATCACTCCCTGGGATCTCTTCAAAAAGATGGTGACACCAAGAA 1855
Db 776 GAGAAGAAGAGAGGAGATCACTCCCTGGGATCTCTTCAAAAAGATGGTGACACCAAGAA 835
Qy 1856 ACCGTTCCGAGACCTCTGAGAGTGCACAGGAGGAGAGCTGGAGAGTCTGAGAGTCAAGAGCC 1915
Db 836 AC-GGTCGGAAGACCTTCTGAGAGTGACAAGGAGGAAGAGCTGGAGAGTCAAGAGCC 894
Qy 1916 CACTTGTCTCTCCACTGATAGCACAGTGTCAAGAAATGCAAGATGAACTCAAACTGTTGG 1975
Db 895 CACTTGTCTCTCCACTGATAGCACAGTGTCAAGAAATGCAAGATGAACTCAAACTGTTGG 954
Qy 1976 TGAGGAACAAAAGCCAGAGAAACAAAGCGTAGGGTGATCTTCACTGTCTTGGGAAGC 2035
Db 955 TGAGGAACAAAAGCCAGAGAAACAAAGCGTAGGGTGATCTTCACTGTCTTGGGAAGC 1014
Qy 2036 ACTGATTTGTGTCGATCATCCAAAGAGAGACCAAGGAAGGCATCTCTTTCAGATGATGA 2095
Db 1015 ACTGATTTGTGTCGATCATCCAAAGAGAGACCAAGGAAGGCATCTCTTTCAT 1071
Qy 2096 AGGAGGCCCAAGGACACTGGGA-GGGGACAGTCAAGAGAGAGGAGGCCAGCAAGACGA 2154
Db 1072 AAGAGGCCCAAGGACACTGGAGGGGGACAGTCAAGAGAGAGGAGGCCAGCAAGACGA 1131
Qy 2155 AAGAAGCCGAAACAGACGCTGTTCTGCCAGCACCCAGGAGGAGACCAGACGCGCAAGAA 2214
Db 1132 AAGAAGCC-GAAACAGACGCTGTTCTGCCAGCACCCAGGAGCAGGACCAAGCGCAAGAA 1190
Qy 2215 GTTCTCACCCGAGCCAGCGGGAAGCCTTCCGAAGGGGAAGGTGTCTCCACTTGGAGT 2274
Db 1191 GTTCTCACCCGAGCCAGCGGGAAGCCTTCCGAAGGGGAAGGTGTCTCCACTTGGAGT 1250
Qy 2275 CATTTAAAGATTAGTCACTCCCAAGAAAAAATCCAAGTCAAAATCTGGAAGAGAAAGCCG 2334
Db 1251 CATTTAAAGATTAGTCACTCCCAAGAAAAAATCCAAGTCAAAATCTGGAAGAGAAAG 1310
Qy 2335 AAGACTCTAGTGTAG-----AGCAGTTGTCTCACTGAGATCGAACCGAGTAGAGAGA 2386
Db 1311 CCGGAAGACTCTAGTTGTAGGAGCAGGTTGTCCACTGAGATCGAACCGGTGAGAGAGA 1370
Qy 2387 ATCTTGGGTTTCCATTAAGAAAATTCATCCCGGACGCGGGAAGAAAGGGCAGACGCGGAA 2446
Db 1371 ATCTTGGGTTTCCATTAAGAAAATTCATCCCGGACGCGGGAAGAAAGGGCAGATGGGAA 1430
```

QY 2447 -GCAAGAACAGCCACTGTGGAGACTAGAGCCAGTGGAGTAAATGAGACACCTTA 2505
Db 1431 GGCAGAAACAAGCCACTGTGGAGACTAGAGCCAGTGGAGTAAATGAGACAGCCTTG 1490
QY 2506 ATGTCCAGCCGTGTGCTGTGTGTGATTAATGAGTGGAGGGAGAAATGTAAG 2565
Db 1491 ATGTCCAGCCGTGTGCTGTGTGTGATTAATGAGTGGAGGGAGAAATGTAAG 1550
QY 2566 CCGAGGGGAATACGAGAGTGCCTCCAGCTGCTGGGGCTGTGTACGTGTCCGAGAGCTCA 2625
Db 1551 CCGAGGGGAATACGAGAGTGCCTCCAGCTGCTGGGGCTGTGTACGTGTCCGAGAGCTCA 1607
QY 2626 GTAAGACTCTGTGTCCAGACTGTGATGTGCGAGTATGATGGAGCAAGGGACATCACCA 1667
Db 1608 GTAAGACTCTGTGTCCAGACTGTGATGTGCGAGTATGATGGAGCAAGGGACATCACCA 1667
QY 2666 GTGTGAGAGAGGGTCTGCTGTGTGATTAATGAGTGGAGTAAATGAGACCTTGAACACA 2745
Db 1668 GTGTGAGAGAGGGTCTGCTGTGTGATTAATGAGTGGAGTAAATGAGACCTTGAACACA 1727
QY 2746 CAGCGGAGAGAGCCATGACCTGTGTAAGAGTCACTGAAAGAACATCATTTGACAGAG 2805
Db 1728 CAGCGGAGAGAGCCATGACCTGTGTAAGAGTCACTGAAAGAACATCATTTGACAGAG 1787
QY 2806 AAACCTCTGTGTCAACCAAGAGTTCAGAGGGTAAAGATGCCATGACGATGTGTCA 2865
Db 1788 AAACCTCTGTGTCAACCAAGAGTTCAGAGGGTAAAGATGCCATGACGATGTGTCA 1847
QY 2866 CCAGTGAAGTGAATTTCACTTCAGAGGCTGTGACAGCCACAGAGACCTCAGAGGGTCTCC 2925
Db 1848 CCAGTGAAGTGAATTTCACTTCAGAGGCTGTGACAGCCACAGAGACCTCAGAGGGTCTCC 1907
QY 2926 GTACTGAAGAGTTCACCAAGAGTCCGAGGGCCGAGAGACACAGACATGTGTCCGAG 2985
Db 1908 GTACTGAAGAGTTCACCAAGAGTCCGAGGGCCGAGAGACACAGACATGTGTCCGAG 1967
QY 2986 TTTTCCAGCTGACTGACTTCCAGACACCAAGAGAGCCACCCCATGTTCAGAGGTGAG 3045
Db 1968 TTTTCCAGCTGACTGACTTCCAGACACCAAGAGAGCCACCCCATGTTCAGAGGTGAG 2027
QY 3046 AGAGTGTGTGTATATTCAGAGAGAGAGAGAGGAGCCAGAGCCATCCCTCAAGCCG 3105
Db 2028 AGAGTGTGTGTATATTCAGAGAGAGAGAGAGGAGCCAGAGCCATCCCTCAAGCCG 2087
QY 3106 TTGCGAGACAGGTGAAAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGACAGAGAG 3165
Db 2088 TTGCGAGACAGGTGAAAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGACAGAGAG 2147
QY 3166 GGTCAAAAGCACTGGAGAGAGTGTGAGAGAGTGAAGAGAGCTCCGAAAGTGTGCTTGG 3225
Db 2148 GGTCAAAAGCACTGGAGAGAGTGTGAGAGAGTGAAGAGAGCTCCGAAAGTGTGCTTGG 2207
QY 3226 AGAAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3285
Db 2208 AGAAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2267
QY 3286 CACAGGGCTCTGAGACTGAGACAGGCTTCAAGAGAGCTTGAAGTCTTGAAGTCAAG 3345
Db 2268 CACAGGGCTCTGAGACTGAGACAGGCTTCAAGAGAGCTTGAAGTCTTGAAGTCAAG 2327
QY 3346 CAGATGTAGACCATGTGCGCACGTGCCAGGTATTAAGTCTCCAGACAGCTGATGAAACAG 3405
Db 2338 CAGATGTAGACCATGTGCGCACGTGCCAGGTATTAAGTCTCCAGACAGCTGATGAAACAG 2387
QY 3406 CCGTGGCCCTGATGTATTCGAAACCTTGAACAGAGTGAACAAATGGAACATCCCT 3465
Db 2388 CCGTGGCCCTGATGTATTCGAAACCTTGAACAGAGTGAACAAATGGAACATCCCT 2447
QY 3466 TAGCAGATTCAAGACTGAGATGGAGACAGAGAGATGAACCAATTGACAGCCAGAGACA 3525
Db 2448 TAGCAGATTCAAGACTGAGATGGAGACAGAGAGATGAACCAATTGACAGCCAGAGACA 2507

QY 3526 GTAAAGCACTGACGCTGTGAGAGGACGACAGAGTCAAGAGAGAGAGGCGCTACTGCTC 3585
Db 2508 GTAAAGCACTGACGCTGTGAGAGGACGACAGAGTCAAGAGAGAGAGGCGCTACTGCTC 2567
QY 3586 AGAAGAGAGAGCCCTTGACACTACCTAATAATGTTCCAGCCACAGAGAGAAACATGGGAG 3645
Db 2568 AGAAGAGAGAGCCCTTGACACTACCTAATAATGTTCCAGCCACAGAGAGAAACATGGGAG 2627
QY 3646 AACCAGAGAGATGTTCTTGAACCTACACAGACAGAGCTTACTGCTGACAGCCGCTCCG 3705
Db 2628 AACCAGAGAGATGTTCTTGAACCTACACAGACAGAGCTTACTGCTGACAGCCGCTCCG 2687
QY 3706 TTTGCGAAAAGACTGAGAGTGGTCAAGAGGTGAGTTGACTGGTTGATGAGAGAAAAG 3765
Db 2688 TTTGCGAAAAGACTGAGAGTGGTCAAGAGGTGAGTTGACTGGTTGATGAGAGAAAAG 2747
QY 3766 TCAAGAGAGACAGAGAGTGTGTTGACACTCTGAGACCCCAACAGTCAAAAGCTGCTGATG 3825
Db 2748 TCAAGAGAGACAGAGAGTGTGTTGACACTCTGAGACCCCAACAGTCAAAAGCTGCTGATG 2807
QY 3826 TGACATATGACATGAAAGTGAATGGAGTGGCGGGGTGTCAAGAAAAGAGATGCTGAAG 3885
Db 2808 TGACATATGACATGAAAGTGAATGGAGTGGCGGGGTGTCAAGAAAAGAGATGCTGAAG 2867
QY 3886 TGCAGAGCTTAAAGCTGTGAGAGAGGAGAGATGGAACGTGACGTTGAAGAGAGAAAAGGG 3945
Db 2868 TGCAGAGCTTAAAGCTGTGAGAGAGGAGAGATGGAACGTGACGTTGAAGAGAGAAAAGGG 2927
QY 3946 AGACAAAGCCAGAGCAAGTGAAGTGAAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAG 4005
Db 2928 AGACAAAGCCAGAGCAAGTGAAGTGAAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAG 2987
QY 4006 AAGAACTCTACGGGAAGCCAGTCTGTGACACTTGAACATGCCCACTGACAGAGGGGAGAG 4065
Db 2988 AAGAACTCTACGGGAAGCCAGTCTGTGACACTTGAACATGCCCACTGACAGAGGGGAGAG 3047
QY 4066 CACTGGGAAGCCTTGGAGAGAGCCCTTCTCTCCAGAGACAAAGACAAAGAGAGGTGCTAG 4125
Db 3048 CACTGGGAAGCCTTGGAGAGAGCCCTTCTCTCCAGAGACAAAGACAAAGAGAGGTGCTAG 3107
QY 4126 AGGTTCAAGTTCAAGCCCTGACACAAACAGTCACTCAACAGCAGAGAGAGAGAGAGAG 4185
Db 3108 AGGTTCAAGTTCAAGCCCTGACACAAACAGTCACTCAACAGCAGAGAGAGAGAGAGAG 3167
QY 4186 TCATPAGAAACGGTGTGTATTTCAAGAGACAGGTGAAGTCCAGAGTGTGTAGTGCACACT 4245
Db 3168 TCATPAGAAACGGTGTGTATTTCAAGAGACAGGTGAAGTCCAGAGTGTGTAGTGCACACT 3227
QY 4246 TATTACCAAGCTGAAGATGCTCTGCAACGGGTGGCCACTGGACTCTTCAGAGTCAAGAG 4305
Db 3228 TATTACCAAGCTGAAGATGCTCTGCAACGGGTGGCCACTGGACTCTTCAGAGTCAAGAG 3287
QY 4306 ACAGGGTACCCCTGGGGCTGAGTCTCAGAGCAGATCCATCCATCATATGATCACTCTG 4365
Db 3288 ACAGGGTACCCCTGGGGCTGAGTCTCAGAGCAGATCCATCCATCATATGATCACTCTG 3347
QY 4366 CTCTGAAAGCACTTACATCTGACCTTCAAGAGAAATAGGCAATCCCAAGAGAGAGC 4425
Db 3348 CTCTGAAAGCACTTACATCTGACCTTCAAGAGAAATAGGCAATCCCAAGAGAGAGC 3407
QY 4426 GATCAGAGAGAGAGACAGAGCAGATGTGTCTGATGTCTGAAGGCAAGAGAGGTACAG 4485
Db 3408 GATCAGAGAGAGAGACAGAGCAGATGTGTCTGATGTCTGAAGGCAAGAGAGGTACAG 3467
QY 4486 CAATCGAAAAGTCTCTCAAGCTGAACTGAGATCTCGGAACTTGAAGTGAAGAGACA 4545
Db 3468 CAATCGAAAAGTCTCTCAAGCTGAACTGAGATCTCGGAACTTGAAGTGAAGAGACA 3527
QY 4546 AGATTGTGCTGAAGCTCAATTCAGACAGCCGTTGAACAGTTGCAAGTGAACAGAGCC 4605
Db 3528 AGATTGTGCTGAAGCTCAATTCAGACAGCCGTTGAACAGTTGCAAGTGAACAGAGCC 3587
QY 4606 CCGAAACTCATGCTTATGATTCACAGAGCCAGAGTTCCTGTCATGCAAGGCTTGAACAGAGG 4665

Db 2028 AGGTGGTGTGCTAGATACAGAGAGAGAGGCCAGAGCCAGGCCATCTCTCAAGCCG 2087
Qy 3106 TTGCAGACAAGGTGAAGAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGCAGAGAACGG 3165
Db 2088 TTGCAGACAAGGTGAAGAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGCAGAGAACGG 2147
Qy 3166 GGTCAAAAGCACTGGAGAGAGTTCAGAGAGTGAAGAGAGTCCGAAGTGTGCTGCTTGG 3225
Db 2148 GGTCAAAAGCACTGGAGAGAGTTCAGAGAGTGAAGAGAGTCCGAAGTGTGCTGCTTGG 2207
Qy 3226 AGAAAGAGAGAGAGCTTTATGCCGAAAGAGCCCGTGCAGGAGAGCTGGAGCTGAGCATCTTG 3285
Db 2208 AGAAAGAGAGAGAGCTTTATGCCGAAAGAGCCCGTGCAGGAGAGCTGGAGCTGAGCATCTTG 2267
Qy 3286 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCAAGG 3345
Db 2268 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCAAGG 2327
Qy 3346 CAGATGTAGACCATGTGCCACAGTGCAGGTTATCAAGCTCCAGCAGCTGATGAAACAGG 3405
Db 2328 CAGATGTAGACCATGTGCCACAGTGCAGGTTATCAAGCTCCAGCAGCTGATGAAACAGG 2387
Qy 3406 CCGTGGCCCTGAGTCACTCCGAAACCTTTGACAGACAGTGAACAAATGGAAGCACTCCCT 3465
Db 2388 CCGTGGCCCTGAGTCACTCCGAAACCTTTGACAGACAGTGAACAAATGGAAGCACTCCCT 2447
Qy 3466 TAGCAGATTGACACACTGCGAGATGGGACACAGCAAGATGAACCCATTGACAGCCAGGACA 3525
Db 2448 TAGCAGATTGACACACTGCGAGATGGGACACAGCAAGATGAACCCATTGACAGCCAGGACA 2507
Qy 3526 GTAAAGCCACTGACAGTGTGAGGAGTCAAGGTTCACAGAGAGAGAGGCGGCTACTGCTC 3585
Db 2508 GTAAAGCCACTGACAGTGTGAGGAGTCAAGGTTCACAGAGAGAGAGGCGGCTACTGCTC 2567
Qy 3586 AGAAAGAGAGAGCTTGCACACTACTTAATATGTTTCAGGCCAGGAGAGACATGGGGAAG 3645
Db 2568 AGAAAGAGAGAGCTTGCACACTACTTAATATGTTTCAGGCCAGGAGAGACATGGGGAAG 2627
Qy 3646 AACAGAGAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGCTGCAGCCGTGCCCG 3705
Db 2628 AACAGAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGCTGCAGCCGTGCCCG 2687
Qy 3706 TTCTGGCAAGACATGAGGTGGGTCAAGAGGTGAGTTGACTGTTGGATGGAGAAAAG 3765
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGTGAGTTGACTGTTGGATGGAGAAAAG 2747
Qy 3766 TCAAAGAGACAGGAGGTGTTGTACACTGTGACCCCAACAGTCAAAAGGCTGCTGATG 3825
Db 2748 TCAAAGAGAACAGGAGGTGTTGTACACTGTGACCCCAACAGTCAAAAGGCTGCTGATG 2807
Qy 3826 TGACATATGACAGTGAAGTGAATGGAGTGGCCGGGTGTCAAGAAAAGGAGAGTACTGAAG 3885
Db 2808 TGACATATGACAGTGAAGTGAATGGAGTGGCCGGGTGTCAAGAAAAGGAGAGTACTGAAG 2867
Qy 3886 TGACAGTCTTAGCTCGAGAGGAGAGATGGAAACTGACTGTTGAAAAGGAGAAAAGGG 3945
Db 2868 TGACAGTCTTAGCTCGAGAGGAGAGATGGAAACTGACTGTTGAAAAGGAGAAAAGGG 2927
Qy 3946 AGACAAAGCCAGACAGTGAAGAGGTGAGCAGGAAAACAGCCGCTCTGAGCATG 4005
Db 2928 AGACAAAGCCAGACAGTGAAGAGGTGAGCAGGAAAACAGCCGCTCTGAGCATG 2987
Qy 4006 AAGGAACTTAGCGGAAGCCAGTCTGACATTTGACATGCCCAGCTCAGAGAGGGGGAAGG 4065
Db 2988 AAGGAACTTAGCGGAAGCCAGTCTGACATTTGACATGCCCAGCTCAGAGAGGGGGAAGG 3047
Qy 4066 CACTGGGAAGCTTTGGAGGAAGCCCTTCTCTCCAGACCAAGAATAAGAGAGGTTGCAATAG 4125
Db 3048 CACTGGGAAGCTTTGGAGGAAGCCCTTCTCTCCAGACCAAGAATAAGAGAGGTTGCAATAG 3107
Qy 4126 AGGTTCAAGTTCAAGGCTTGACACAAACAGTCACTCAAAACAGACAGAGAGGCTGTGAAAAGG 4185

Db 3108 AGGTTCAAGTTCAAAGCTGGACACAAACAGTCACTCAAAACAGCAGAGAGCTGTGAAAAGG 3167
Qy 4186 TCATAGAAAACGGTTGTGATTTTCAGAGACAGGTGAAAAGTCCAGAGTGTGTAGGTGCACACT 4245
Db 3168 TCATAGAAAACGGTTGTGATTTTCAGAGACAGGTGAAAAGTCCAGAGTGTGTAGGTGCACACT 3227
Qy 4246 TATTACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGAG 4305
Db 3228 TATTACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGAG 3287
Qy 4306 ACACGGTACCCTTGGGCGCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACTCTG 4365
Db 3288 ACACGGTACCCTTGGGCGCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACTCTG 3347
Qy 4366 CTCTGAAAGCACCTCATCTGACCTACAAGGAGAAATTAAGCGCATCCAGAGAGAGC 4425
Db 3348 CTCTGAAAGCACCTCATCTGACCTACAAGGAGAAATTAAGCGCATCCAGAGAGAGC 3407
Qy 4426 GATCAGAGGAGAGGACAAAGCCAGATGCTGCTGATGCTGACGGCAAGGAGATGACAG 4485
Db 3408 GATCAGAGGAGAGGACAAAGCCAGATGCTGCTGATGCTGACGGCAAGGAGATGACAG 3467
Qy 4486 CAATCGAAAAGTCTCAAGGCTGAACCTGAGATCTCTGGAACCTTGAGAGTAAAGAGCAACA 4545
Db 3468 CAATCGAAAAGTCTCAAGGCTGAACCTGAGATCTCTGGAACCTTGAGAGTAAAGAGCAACA 3527
Qy 4546 AGATTGTGCTGAACGTCATTTCAGACAGCCGTTGACCAAGTTTCGCACGTACAGAAACAGCCC 4605
Db 3528 AGATTGTGCTGAACGTCATTTCAGACAGCCGTTGACCAAGTTTCGCACGTACAGAAACAGCCC 3587
Qy 4606 CCGAAAATCATGCTTATGATTCACAGACCAGGTTCTGTCATGTCAGGCTTGACAGAGGG 4665
Db 3588 CCGAAAATCATGCTTATGATTCACAGACCAGGTTCTGTCGAATGCGCTTGACAGCAGGAGG 3647
Qy 4666 AGCCACACAGATCTCGACAAAATTAAGAGATCCAGAGATGAAACACCCAGTCCGCGCAGC 4725
Db 3648 AGCCACACAGATCTCGACAAAATTAAGAGTTCGAAGATGAAACACCCAGTCCGCGCAGC 3707
Qy 4726 CCAGAGAGAGACTTGGAAAGCTCTGACCGTTCTGAGAGGATGGGCTCAGCCTCGGAAATGCT 4785
Db 3708 CCAGAGAGAGACTTGGAAAGCTCTGACCGTTCTGAGAGGATGG--CTCAGCTCGGAAATGCT 3765
Qy 4786 TGCCGCGCTTGCAGTTGAAAAGCCCGGTGTCAAAGTAAAGCATTTGAGAAAGTGTGCTCTCTCA 4845
Db 3766 TGCCGCGCTTGCAGTTGAAAAGCCCGGTGTCAAAGTAAAGCATTTGAGAAAGTGTGCTCTCTCA 3825
Qy 4846 ACCCAAGATCCAAAAGGAGCATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4905
Db 3826 ACCCAAGAT--CAAAAAGGAGCATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 3884
Qy 4906 CAGAGGCCAGTGTCTCTGGAAAACCTAAACAAAAGAAATCCCCAGACACCAACGGACCAAAAGC 4965
Db 3885 CAGAGGC--AGTGTCTGGNAACCTAAACNAAGAAATCCCCAGACACCAACGGACCAAAAGC 3941
Qy 4966 TAAACCGAGGAGGCGATGCCCCAAAAGTTGAGTCCAGGAAGAAAGAAATGCTTCAACAAAGT 5025
Db 3942 TAAACCGAGGAGC-ATGCCCCAAAAGTTGAGTCCAGGAAGAAAGAAATGCTTCAACAAAGT 4000
Qy 5026 CAGTCAAGAGAACAAAGGCCAGGACAGAGAGACCTGCGAGGAGCCCAAGGGAGACCTGG 5085
Db 4001 CAGTCAAGAGAACAAAGGCCAGGACAGAGAGACCTGCGAGGAGCCCAAGGGAGACCTGG 4060
Qy 5086 CAGAATCTTAAGATGTTAGTTGCTCATTTGATCATCTGTAAGACCAAGATGTGAAAACAAG 5145
Db 4061 CAGAATCTTAAGATGTTAGTTGCTCATTTGATCATCTGTAAGACCAAGATGTGAAAACAAG 4120
Qy 5146 TCACAGAACAAAGATGCTGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATC 5205
Db 4121 TCACAGAACAAAGATGCTGCTGTTGGGACCTTG--GACCAAGATTTTCAGAGCCCATGAGATC 4179
Qy 5206 CAGAGAGCAGGCGGCTCCAAATGATTTCCACCCAGTGAAGCACCCCGACAAATTCCTGAGGCT 5265
Db 4180 CAGAGAGCAGGCGGCTCCAAATGATTTCCACCCAGTGAAGCACCCCGACAAATTCCTGAGGCT 4239

964 CTGCCGAGAGAAAGGAGCAAGAGGAGAGAAAGTAGACGAGCAAGAAAGGAAAGA 1023
1070 CTTCCGAGAGAGAAAGGAGCAAGAGGCTGAAAGTAGACGAGCAAGAGGAAAGA 1129
1024 CAGAGCCAGCC---TCGAGAGAGAGAGGAGCGGAGAGAGACACAGAGCCAGGTTGT 1080
1130 CAGAGCCAGCCAGCCAGGAGAGAGGAGCTGAGAGAGGAGAGAGAGGAGGAGGTTGT 1189
1081 CAGCAGACTAGAGAAAGGTGGAGCTGCTTTGGAGAGAGAGAGGAGGAGGAGGAT 1140
1190 CAGCGGACTAGAGAGGAGGAGTGGCTTTGGAGAGAGAGGAGGAGGAGGAGGAT 1249
1141 CGTCAGAGAGAGAGGAGTGGCTTTGGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1200
1250 TGTC---GGAGAGTGTCTCTTTGGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1306
1201 ACCAAGAGTGTGTGAGAGGAGTCCAGCTGAGAGAGGAGGAGGAGGAGGAGGAGG 1260
1307 ACCAAGAGTGTGTGAGAGGAGTCCAGCTGAGAGAGGAGGAGGAGGAGGAGGAGG 1364
1261 GAGGAGAGAGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
1365 -AGGAGAGAGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1423
1321 CTGAGAACTGGCTGAGCCAGGAGGAGTCCAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
1424 CTGAGAACTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1483
1381 TGAAGAGCAGAGAGAGTGTGTCTCTGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1440
1484 TGAAGACCAAGAGAGTATGGCTCTCTGGGAGTGAACATCTAGCTGAGAGTCTAG 1543
1441 CTGAGAGAGAGAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
1544 CTGAGAGAGAGAGTCTTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1603
1501 CCTCTAGGAGAGAGTCAAGGAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
1604 CCTCTAGGAGAGAGTCAAGGAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1663
1561 GCTTAAAGAGAGTCTCTGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1617
1664 GCTTAAAGAGAGTCTCTGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1723
1618 AAGAGCTGGAGAGATCAACACATTCACACCGAGTCCCGAGAGAGGAGGAGGAGGAG 1677
1724 AAGAGCCAGGAGAGATCAACACATTCACACCGAGTCCCGAGAGAGGAGGAGGAGG 1783
1678 AAGGAGAGAGTCTGCGTCTGCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1737
1784 AAGGAGAGAGTCTGCGTCTGCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1843
1738 GCGCGCTGGAGAGACCCAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1797
1844 GCGCATCGGAGAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1903
1798 AAGAGAGAGAGGAGTCACTCCCTGGGAGTCTTCAAAAGAGAGGAGGAGGAGGAGGAG 1857
1904 AAGAAAGGAGAGGAGTCACTCCCTGGGAGTCTTCAAAAGAGAGGAGGAGGAGGAGG 1963
1858 GGGTCCGAGAGAGTCTGAGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1917
1964 GGGTCCGAGAGAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2023
1918 CTTGTCTCTCAGTATGAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1977
2024 CTTGTCTCTCAGGAGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2083
1978 AGGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2037
2084 AGGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2143
2038 TGATTGTGTGGAGTATCTCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2097

2144 TGATTGTGTGTGGATCGTCCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2203
2098 GAGGCCCAAGGACACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2157
2204 GAGGCCCAAGGACACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2263
2158 AAGCCGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2217
2264 AAG-----CAGATGTCTTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2317
2218 CTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2277
2318 CTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2377
2278 TTTAAAGATTAGTCACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2337
2378 TTTAAAGATTAGTCACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2437
2338 ACTCTAGTGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2397
2438 ACTCCGAGTGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2497
2398 CCATTAAAGATTATCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2457
2498 CCATTAAAGATTATCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2557
2458 CCATTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2517
2558 CCGCGTGTGAAGACTCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2617
2518 TCGTGCCTCTGTCTGATGATTAATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2568
2618 TTGTGCCTCTGTCTGATGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2677
2569 AGGGAGATACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2628
2678 AGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2737
2629 AGACTCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2688
2738 AGACTCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2797
2689 TCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2748
2798 CCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2857
2749 CCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2808
2858 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2917
2809 CTCTGTGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2868
2918 CTCTGTGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2977
2869 GTCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2928
2978 GTCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3037
2929 CTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2988
3038 CTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3097
2989 CCAGAGTGTGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3048
3098 CCAGAGTGTGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3157
3049 GT 3108
3158 GT 3217
3109 CAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3168

Db 3218 CAGACAAAGTAAAGAGACTCCAGGTGCTGCAACCCAGACTTGACAGAGAGGAC 3277
 Qy 3169 CAAAAGCACTGAGAGAGTTGAGAGGTAGAGAGAGACTCCGAAATGCTGCTTGAGAGA 3228
 Db 3278 CGAAAGCACTGAGAGAGGTGAGAGGTAGAGAGAGACTCCGAGGTGCTGAGCTACCGAGA 3337
 Qy 3229 AAGAGAAAGAGCTTATGCGCAAAAGAGACCCGTGCAAGAAAGCTGAGACTGACATCTTTGAC 3288
 Db 3338 AAGAGAAAGAGTGTGTGCGGAGAGACCCGTGCAAGAAAGCTGAAACTGAGCATCTTTGAC 3397
 Qy 3289 AAGGCTGAGAGCTGAGAGAGCTACTCCAGAGAGCTTGAAGTTCTTGAAGTACAGGAC 3348
 Db 3398 AAGGCTCGAGAGCTTACAGGTTACCCAGAGAGCTTGAAGTTCTTGAAGTACAGAGG 3457
 Qy 3349 ATGTAGACCATGTGCGCAAGTGCAGGTTATCAAGCTCCAGAGCTTATGAAACAGGCCG 3408
 Db 3458 ATGTAGACCGTGCACACATGTCAGGTTATCAAGCACAGAGCTGATGAAACAGGCCG 3517
 Qy 3409 TGGCCCCCTGAGTCACTCCGAAACCTTGAAGAGAGTGAACAAATGGAAGCACTCCCTTAG 3468
 Db 3518 TGGCCCCCTGAGTCACTGAAACCTTGAAGAGAGTGAACAAATGGAAGTACTCCCTTG 3577
 Qy 3469 CAGATTGAGACACTGAGAGTGGAGACAGCAAGATGAAACCTTGAAGAGCCAGAGAGTA 3528
 Db 3578 CAGATTGAGACACTGAAACCGGAGACAGCAAGAGAGAGCTTGAAGAGCCAGAGAGTA 3637
 Qy 3529 AAGCCACTGAGAGCTGTGAGGAGTCAAGAGTCAAGAAAGAGAGCGGCTACTGCTCAGA 3588
 Db 3638 ATGCCATTGCGCGCTCAAGAGTCAAGAGTCACTGAAGAGAGGCAAGCTGCTGCTCAGA 3697
 Qy 3589 AAGAGAGAGCTTCGACACTACTAATATATGTTCCAGCCCAAGAAACATCTGGGAGAAAG 3648
 Db 3698 CGAGAGGAGCTTCAACACCATC --- TAGTTTTCAGCCCAAGAAACATCAAGGAGAAAC 3754
 Qy 3649 CAGAGAGAGATGTTGTTGAACCTACAGCAAGAGTTTACTGCTGAGAGCGGCTGCTTC 3708
 Db 3755 CAGAGAGAGATGTTGTTGAACCTACAGCAAGAGTTTACTGCTGAGAGCGGCTGCTTTC 3811
 Qy 3709 TGGCAAGAGCTGAGAGTGGGTCAAGAGGAGTGAAGTTGACTGTTGAGTGAAGAAAGTCA 3768
 Db 3812 TGGCAAGAGCTGAGAGTGGGTCAAGAGGAGTGAAGTTGACTGTTGAGTGAAGAAAGTCA 3871
 Qy 3769 AAG-----AAGAAACAGAGAGTGTGTTGACTCTGGAACCCACACTCAAA 3813
 Db 3872 AAGAGCGAGAGTGTGTTAAGAACTGAGAGTGCCTGTGACACTGAAACACTCAAA 3931
 Qy 3814 AAGGCTGAGATGATGACATATGACAGTGAAGTGAAGAGTGGCCGGGTGTGAGAAAG 3873
 Db 3932 AAGCTGCTGACTTGAACGTGACGTGACAGTGAAGTGAAGTGGCCAGATGTCAAGAAACTG 3991
 Qy 3874 AAGTACTGAAGTGAAGTCTTAACTTGAAGAGAGAGAGATGAAACTGAAGTTGAA 3933
 Db 3992 AAGTACTGAAGATCAGATATTAGCCCGGAGAAAGAGATGGAACCGAGCTTGAA 4051
 Qy 3934 AAGAGAAAGAGAGACAAAGCCAGAGCAAGTGAAGAGAGTGAAGAGAGAAACAGCCG 3993
 Db 4052 AAGAGAAAGAGAGACAAAGCAAGAGCAAGCCAGTGAAGAAACATGAGAGAGAAACAGCTG 4111
 Qy 3994 CTCCTGAGATGAAGAGACCTTACGGGAGAGCCAGTCTGACACTTGAAGAGCCAGCTCAG 4053
 Db 4112 CTCCTGAGATGAAGAGACCTTACGGGAGAGCCAGTCTGACAGTGAAGTCTCACTCAG 4171
 Qy 4054 AAGAGGGAGAGAGCACTGGAGAGCTTGAAGAGAGCCCTTCTCCAGACCAAGACAAAG 4113
 Db 4172 AAGAGGGAGAGAGCACTGGAGAGCTTGAAGAGAGCCCTTCTCCAGACCAAGACAAAG 4231
 Qy 4114 CAGGTGATAGAGGTTCAGTTCAAGCTTGAAGACCAAGTCACTCAACAGAGAGAG 4173
 Db 4232 CAGATTGATAGAGGTTCAGTTCAAGCTTGAAGACCAAGTCACTCAACAGAGAGAG 4291
 Qy 4174 CTGTGAGAAAGGTTCATAGAAAGGTGTTGATTTGAGAGCAAGTGAAGAGTCCAGAGTGG 4233
 Db 4292 CTGTGAGAAAGGTTCATAGAAAGGTGTTGATTTGAGAGTGAAGAGTTCAGAGTGG 4351

Qy 4234 TAGGTGCACTATTATACAGCTGAGAGATCTTGCAGAGGGTGGCACTGCACTCTTC 4293
 Db 4352 CAGGTGGCAATCATTAACAGCTGAGAGAGCTTCCGAAACCGGTGGCTACCGGACTCTTC 4411
 Qy 4294 AAGATGAGAGAGAGACCGGTGACCCCTGGGGCTGAGTCTCAGGCGAGATTCATCCCAATCA 4353
 Db 4412 AAGATGAGAGAGAGACCGGTGACCCCGGGGCTTGAAGTCTCAGGCGAGATTCATCCCAATCA 4471
 Qy 4354 TAGTAACTCTGCTCTGAAAGACCCCTACATCTGACCTACAGAGAGAAATTAAGCGCAT 4413
 Db 4472 TAGTAACTCTGCTCTGAAAGACCTTCAATCTGACCTTCAAGAGAGAGTGGCGCAT 4531
 Qy 4414 CCGAGAGAGAGAGATCAGAGAGAGAGACCAAGATGCTGCTGCTGAGTCTGACGCA 4473
 Db 4532 CCGAGAGAGAGAGATCAGATGAGATGAGATCAAGAGATGCTGCTGAGTCTGACGCA 4591
 Qy 4474 AAGAGAGTACAGCAATGGAATAAGTCTCAAGAGTGAAGCTGAGATCCTGGAACCTTGAGA 4533
 Db 4592 AAGAGAGTGCAGAGAGAGAGAAATCTCAGAGGCTGGAACCTGAGATCTTGAGACTTGAGA 4651
 Qy 4534 GTAAGAGCAACAGATTTGTGCTGAACGTCAATTCAAGACAGCCGTTGAACCAATTGACAGTA 4593
 Db 4652 GTAAGAGCAATTAAGATTTGTCCAGAGTGTCAATCCAGAGCCGTGACCAAGTTGACAGTA 4711
 Qy 4594 CAGAAACAGCCCCGAAACTGATGCTTATGATTCAGACAGCCAGTTCCTG-CATGCAAG 4652
 Db 4712 CAGAAACAGCCCCGAAACCCACGCTTCTGATTTACAGAAATCAGATGCTGATGAGAGG 4771
 Qy 4653 CTGACAGCAGAGAGGCCCAAGATGCTGAGCAAAATGAAAG-----AT 4697
 Db 4772 CTGACAGCAGAGAGAGCAACAGATGCTGAGCAAAATGAAAGAGCACTTCAAGTCTGCC 4831
 Qy 4698 GCCAAGTGAATAACCCAGATGCGCAGAGCCAGAGAGACTTGAAGTCTGAGCCGTTCTG 4757
 Db 4832 CCGAAGATGAACACTCAGATGCGGTAGCCAGAGAGAGACTTGGGTTTCTGATGTTCTG 4891
 Qy 4758 GAGCATGAGGCTCAGCTCGGAAATGCTTGCAGGCTTGAAGTGAAGCCCGGTGCA 4817
 Db 4892 AAGCATGAGCAAGCTTCAAAATGATCAACAGCTTGAAGTGAAGTCCAGTGTCA 4951
 Qy 4818 AAGTAAAGATTAAGAGAGTGTCTCTCAACCCAAAGATCCAAAGAGAGAGATGCTGAT 4877
 Db 4952 AAGAAAGTGTGAAGAGTGTCTCTTCAAGTCAAGAT-GAAAGAGAGATGCTGATGAC 5010
 Qy 4878 GGCCTCAGCTTCAAGAGTTAGCCAGAGCAAGGCAAGTGTCTTGAAGACCTAACAA 4937
 Db 5011 GGCCTCAGCAACAAAGTTAGCCAGAGCAAGGCGAGATGCTTGGAAATCTAACAA 5070
 Qy 4938 GAATGCCAGACCAACAGAGCAAGCTTAACGAGAGAGGCGATCCCCCAAAAGTTGAG 4997
 Db 5071 GAGTCCCGAGACCAACAGAGCAAGCTTAACGAGAGAGAGATGCTTGAAGAGAA 5130
 Qy 4998 GTCCAGAGAGAGAAATGTCTAACAGATCACTCAAGAGAGCAAGGCCAGAGAGAGAG 5057
 Db 5131 AT-----GAACAGAGCCAGAGAGAGAG 5154
 Qy 5058 ---GACTGAGAGCCAAAGAGAGACTGAGAGATCTTAAGATGTTAGTGTGCTATTG 5114
 Db 5155 GAGAGCTTACAGAGAGCCAAAGAGAGACTGAGAGATCTTAAGAGTGAAGTGTGCTCAT- 5212
 Qy 5115 TACATCTGTAAGCCAGATGTGAAGAAACAGTCAACAGAAAGATGCTGCTGTTGGAGCC 5174
 Db 5213 -----TGTAAATCTGAAGTGAAGTGAAGTCAACAGAAAGATGCTGCTGTTGGAGCC 5266
 Qy 5175 TTGAGACCAAGATTTCAAGAGCCATGAGATCCAGAGAGAGAGGCGCTTCAATGATTTCCA 5234
 Db 5267 TTGAGACCAAGATTTCAAGAGCCCTTGAAGTGAAGAGCAAGAGCCGCTCAATGATTTCAA 5326
 Qy 5235 CCGAGTACAGACCCCGCAATTTCTGAAGCTTCAATCGGAGCTGAGAGCACTAACTTT 5294
 Db 5327 CCTGAGAGAGACCCCGCAATCTGAGAGCTGCAATCGGAGCTGAGAGCACTAACTTT 5386

Db 1035 ACTAGTCCCGTGCACGATGTAACAGATCACTTCAAAAATTTCTTCACTCAAGTTGG 1094
Qy 909 GCCGGCTGCGCAGAGAACAGCAGCTTCAAGAAATCAAAAAGAGATGATCTGAAAATGCGC 968
Db 1095 GCCGGCTGCGCAGAGAACAGCAGCTTCAAGAAATGAAAGAGATGAAAGTGAACCTTCA 1154
Qy 969 GAGAAAGAAAG 1028
Db 1155 GAGAAAGAAAG 1214
Qy 1029 CCAGCTCGAG 1061
Db 1215 GTTGGCTCCGAG 1274
Qy 1062 ACAGACAG 1121
Db 1275 GCCACAG 1334
Qy 1122 GTTGGTACCTGAG 1181
Db 1335 GTCACTGGCTCCAG 1394
Qy 1182 GATGAG 1238
Db 1395 GATGAG 1454
Qy 1239 GAG 1298
Db 1455 GAG 1490
Qy 1299 GGAACAG 1358
Db 1491 GAAACAG 1550
Qy 1359 GGTGAGCTGCTGAG 1418
Db 1551 GCCGAACTGCGCAG 1610
Qy 1419 ACTCACTGACAG 1478
Db 1611 AACAG 1670
Qy 1479 GTCACTGAG 1538
Db 1671 GTGAG 1730
Qy 1539 AAGAACTCTTCACTGAG 1598
Db 1731 AAGAACTCTTCACTGAG 1790
Qy 1599 CGAG 1658
Db 1791 AAG 1847
Qy 1659 GAG 1718
Db 1848 GAG 1907
Qy 1719 ACCAGCTGCTGAG 1778
Db 1908 ATCACTGCTGAG 1967
Qy 1779 ACTACTTCCGATGAG 1838
Db 1968 GCTACTTCCGATGAG 2027
Qy 1839 ATGAGTGAACCCAG 1898
Db 2028 ATGAGTGAACCCAG 2087
Qy 1899 GAGAAAGTGAAG 1958

Db 2088 GACAAAGTCAAG 2147
Qy 1959 GAATGTCAAACTGTTGGTGAAG 2018
Db 2148 GAAATGAAG 2207
Qy 2019 TCAGTGTCTTGGAG 2078
Db 2208 TCAGTGTCTTGGAG 2267
Qy 2079 TCTCTTCAATGATGAAG 2138
Db 2268 TCTCTTCAATGATGAAG 2327
Qy 2139 GAGGTCAG 2198
Db 2328 GAGGTCAG 2387
Qy 2199 GACCAAGGCGAAG 2258
Db 2388 GATCCAGGCGAAG 2447
Qy 2259 GTTCTCACTTGGAG 2318
Db 2448 GTTCTCACTTGGAG 2507
Qy 2319 CTGGAAG 2366
Db 2508 CTGGAAG 2567
Qy 2367 ATCGAACGAGTGAAG 2426
Db 2568 ACTGAACCGGTAAG 2627
Qy 2427 AAGAAAGGCGAAG 2486
Db 2628 AAGAAAGGCGAAG 2687
Qy 2487 ATAAATGAG 2546
Db 2688 GCACACGAG 2747
Qy 2547 GAG 2597
Db 2748 GAAAGGAG 2807
Qy 2598 GGGGCTGTGATGAG 2657
Db 2808 GCAGCACGAG 2867
Qy 2658 GTCAATGATGAG 2717
Db 2868 GTGCTGACGAG 2927
Qy 2718 GCTTCGTAACAG 2777
Db 2928 GCTTCGTAACAG 2987
Qy 2778 GTCACTGAAAAAG 2834
Db 2988 GTATTGAAAAAG 3047
Qy 2835 GAGGTAAG 2894
Db 3048 GAGAACAG 3107
Qy 2895 GTGACAGCAAG 2954
Db 3108 GTGACAGCTGCAAG 3167
Qy 2955 GCCGAAG 3014
Db 3168 GCTGAAG 3227

Db	5382	CCAAACCAAAAGAGAGAGATGCCAGGAAGTAAATTGCAGGAAGAAAGAAAGTGCAC	5441
Oy	5019	ACCAAGTCAGTCAAGAGAGACAAGGCCGAGGAGAGAGACCTGCAGAGACCAGAGGA	5078
Db	5442	AGGAAATCAGATTAAGGAGATCACACCCCAAGCAGAGGAGTTACAGAAAACAAGAGAGA	5501
Oy	5079	GACCTGGCAGATCTTAAGTGTAACTG-----CTCATGTAC	5117
Db	5502	GAATCTGCAAAAGTCAGAACTTACAGAAATTTAAAAATCATTGACGTAACTCATTTGCT	5561
Oy	5118	ATCTGTAGAACCAGAAATGTGAAAAACAAGTCACAGAACAA----GATCGTGTGTGGGAC	5173
Db	5562	GTTTGGAAAGCCAGAAATGTGAAACAAAGTAAAGTAAAGAAATGAAATGCTGCTGTGAGA-	5620
Oy	5174	CTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGACAGAGGCCGTCAATGATTTCC	5233
Db	5621	CTGAAGACCAAGTATTTTACAG-ACTTTGAGAAATTTGAGAGACAGGACATCACTGATCTCA	5679
Oy	5234	ACCCAGTAGAGACACCCGACAAATTCGAGCTTCATCGGAGCTTAGACCAGCTAACATT	5293
Db	5680	TTTCTAAGAGAGC-CCCTGACAATCTCGAGGCTTCATCAGAGACTTAGACCATTTPAACATT	5738
Oy	5294	TCCTCGTTTCAAGCTGCCCTTTGATTTGGCCCTGTATGCCGTCGATATTTCTAACATT	5353
Db	5739	TCCTCTTTCAGAGCCAACTCAAAATTTCCCTGTAAAC-----	5778
Oy	5354	TCCTCGTTTCAAGACTCCCTTTGATTTGGCCCTGTATGCCGTCGATATTTCCGATTTTA	5413
Db	5779	-----CATATAAATTCGATTTA	5796
Oy	5414	AGGTCCTGCGTTCTCAACTCGGAAC-CAAATTCGCATATCTTAAGTTCCACTTCTCAAACT	5472
Db	5797	AGGTCCTAAATTTCTTAACCTCGGAACCTGGAATGCGCAATCTCAATTCGCTTCGAAACT	5856
Oy	5473	GGAGATCCCTCTTAAGTAAATTAATTAAGTAAATGATTAAGTAAAGCCCTCCGTCACAT	5532
Db	5857	GGAGATCAATCTTTAACATTTAATTAATTAAGTAAAGTAAAG---TCCTCTGATCTAAT	5913
Oy	5533	TGTAATTTTTTTCTAAAGTTT-----AAGCAATGCTTTTGTATTA	5575
Db	5914	TGTAATTTTTTTCTTAATGTTTAAGTAATGTCAGAAATCAATGCTTTTGTATCA	5973
Oy	5576	TGCATATATTAACGGGTGTCAGGCCATAGCGACGCTTTGAAAAAGCTCAAGCTCAACTG	5635
Db	5974	CACAGTATATGATGGGCAATGTGCAATAGTGCAGGCTTGAGGAGCTTTAAGCTCAGTTA	6033
Oy	5636	TAACTGCAGCAAAACAGATACATTC-----CTGGCAAGAAAGACCAAGTCTTTT	5686
Db	6034	TATAACCCACAAAAACAAGCCCTCCATGATGTAACTTCGTATCAAGGTACAATTCCT	6093
Oy	5687	TAAAGTTTACGATGCTTAAGATCTGTGGGCTTCAGTCCCTGAAAGTGTTGTTTCTCT	5746
Db	6094	TAAATTTACCTAAAGATTTAGAGTCCATATTTAGTGTACTGTAAATTTGTCACTTTCTCT	6153
Oy	5747	ATGCACACGAGCTCAGAAATTAAAAAACCATTGTAACATCCAGAGATGCCAATATT	5806
Db	6154	ATTACACGAGATGTCTAAACCTAAAGAAATTTGAACAATACAGAAATGTTCTAATGTC	6213
Oy	5807	ACCATGATTTTTTCCCCCTTTTGTCTAATTCAGTCCAGGTTGGAAAAAGTCTCTCTG	5866
Db	6214	ATTGGAAATTT-----TTCTTCTAACCCAGTGGAGGTTGAAAGAAATTAATTCCT	6266
Oy	5867	TGTCAAG-----TTAAGCCCTGTCTTAAGATGATGATGACAAATGAGT	5909
Db	6267	GGTACCAATTAACCTTAACATCTTTTCTCACTGTTATGTTGTTTGGACCAATTAAGT	6326
Oy	5910	GTGCTTAAGGCATAGAC-----TGTTTCTAATGCAGAGAAATCTGT	5953
Db	6327	GTGCTTAATCTCGAGGCAAAATAGTGAATATGTTTAAATATGTTAAGAAAAAAGAAATGT	6386
Oy	5954	TGTACGTTTTTTTGAATGTACTCTTCTATGCTGAGCCGAATTCATATGCAATCGAAGTG	6013

Accession	Sequence	Position
Db	TGTAAAG--TTTTCGATTTCACTCTTAATATGCTGACCTGCATTCACACATGACATGAATA	64444
Oy	AGTCCTGTCTTTACACAGATGATTTTGAATAATACATGACAGTTTCTCTGTATATATCTG	60733
Db	AGTACAGTTCTTTACAAATGGATTTTGTATATATCTGCATTTGTTTGTGCATTAATTTG	65040
Oy	TGCCCTCTCTTTAAGAACATGTGCATTAATGTTCCCTTGGATTAATTTGTATTTGACAA	61333
Db	TGCCATTTCTTTAAGAACATTTGTGCACACATTCATTTTGGATTAAGTTGTGATTTGACGA	65644
Oy	CTGATTTAATATAA	61447
Db	CTGATTTAATATAA	6578

```

RESULT 13
US-60-278-232-10143
: Sequence 10143, Application US/60278232
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preeti
: TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
: TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
: TITLE OF INVENTION: Polymorphisms Identified Thereby
: FILE REFERENCE: GX-0011 P
: CURRENT APPLICATION NUMBER: US/60/278,232
: CURRENT FILING DATE: 2001-03-30
: NUMBER OF SEQ ID NOS: 12,557
: SOFTWARE: PERL Program
: SEQ ID NO 10143
: LENGTH: 6603
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No: 445786.12
US-60-278-232-10143

```

Query Match	Similarity	35.4%	Score 2178;	DB 71;	Length 6603;
Best Local	Similarity	64.4%	Pred. No. 0;		
Matches 4222;	Conservative	0;	Mismatches 1795;	Indels 537;	Gaps 43;
3	GGCCGGCGGGAATGAAAGACCACTGAGCCATGAGGCGCAGGCAATTCCACCGAGACGCG	62			
Db	GGCTAGGCGCGGGAAGAAATGCGGAGAGGACATATGGGCGCGGAGGCTCCACCGAGCAGCGC	222			
63	AGCCCGAGCAGCCG---CGGGGAGGCCACATGCCGAGCGACTGTGTCTAGTGGCCAT	119			
Db	AGCCCGAGCAGCACCAGCCCGCAGGGGAGGCTCCACGCGGCTGACCCGAGGCCAGCGCGGCG	282			
120	GGGGCCCGCAGCTGAAAGCCCTGGGGA-----GCAGCTGGAGACCCCGCCACGCGGAC	170			
Db	GGGCTCTCGGCGGAGGCGGCGCCGACACACCCGCGGACCCCGCATCGCTGCTGGAC	342			
171	CCGCGCACCAGGCTCCACAGAAAGATGGCCAGCTGTCTTCTGTCAACGCGTACGTGA	230			
Db	CCGCGCACCAAGTCTCTACAGAAATGTGTCAGCTGTCCACCATCAATGGCGTATGAG	402			
231	CAAGGAGATGTCATGTCCCAAGAGGAAAACCAGAGGG-----	268			
Db	CAAGATGAGCTCAGCCTCCAGAGGGGTGACTTAAATGGGCAAGAAAGAGCCTTGAACGGT	462			
269	-----GCAGGAGAAAGATGCTTATGATGAGGATTTGGACAGCA	308			
Db	CAAGGAGCCCTAAACAGCCAGAGGAGGAAGAAGATCATTTGTACAGAGGTTGGACAGAGA	522			
309	GAGTCAGAAAGATGTGAGAGAAAAAGCCAGATTGAAGAAATGGCGGCCAATCTCACAGCT	368			
Db	GACTTGAAGATGTGAGCAAAAGAGACTCCGATTAAGAGATGGCTCTTAAGTCAGCGGTT	582			
369	GTTGAAGATATCAAAAGATGGCGAGGAGGAGCATCGAAATATATTGAACGATCCCT	428			

Db 583 GTTCACGACATCAGATGATGGCAGGAGGAGACCCCGAAATAATCGAACGATTCCT 642
Qy 429 GCTTCAGAAACAATGTGGAAGAAATGGTACAGCCTGCTGAGTCCCGAGGTAAATGATGTT 488
Db 643 TCTTCAGAAAGCAATTTAGAGAGCTAAACAACCCACTGAGTCCCGAGGTAAATGATATT 702
Qy 489 GGCTTCAAGAAAGTATTTAAATTTGTTGGTTTAAATTTCAAGTGAAGAGGATAAAAT 548
Db 703 GGATTTAAGAAAGGTGTTTAAAGTTTGTGGCTTTAAATTTCACTGTGAAGAAAGGATAAGACA 762
Qy 549 GAAAGTCAGATACGTCCCAACTACTCACTGTCAAGAGGATGAAGCGAAGGGCCAGAA 608
Db 763 GAGAGCCTGACACTGTCAGCTACTCACTGTGAAGAAAGTGAAGGGGAGGAGCAGCA 822
Qy 609 GCCTCTGTGGAGCTGGAGACCAACAGAGCCAGTGTGGAGACTGCGGTGGAGAGTCA 668
Db 823 -----GGGGGTGGCGACCAACAGAGCCCGAGCCTTGGGGCT-----GGAGAAGCA 867
Qy 669 GCATCCAAAGAAAGTGAAGCTGAAGCAATCCACAGAGAGCAAGAGGACCCCTGAAGCAA 728
Db 868 GCATCCAAAGAAAGCGAACCACAAACAATCTACAGAGAAACCCGAGAGACCCCTGAAGCGT 927
Qy 729 GAAACAGAGCAGCAGAAATCCCTCTTCAAGCCGAATCTGATCAAGCGGCTGAGGAAGAA 788
Db 928 GAGCAAGCCAGCGAGAAATTTCTCCCGCAGCCGAATCTGGCCAAAGCA--GTGGAGGAA 984
Qy 789 GCCAAAGATGAAGAGAGAAAGAAACAAGAGAAAGAGCCCAAGTCCCAAGAAATCCCG 848
Db 985 TGCAAGAGGAGAGAGAGAAACAAGAAAGAAACCTAGCAAGTCTGCAGAAATCTCCG 1044
Qy 849 AGCAGCCAGTCAACAGTGAAGACACATCTTCTTCAAGAGTTCTTCACTCAGGTGG 908
Db 1045 ACTAGTCCCGTGACCGAGTGAACAGGATCAACCTTCAAAAATTTCTTCACTCAAGGTGG 1104
Qy 909 GCCGCTGGCCAAAGAGACCAAGCTTCAAGAAATCAAAAGAGGATGATCTGGAACCTGCC 968
Db 1105 GCCGCTGGCCAAAGAGACCAAGTTTCAGAGCGGAGAGGAGATGAAGTGAAGCTTCA 1164
Qy 969 GAGAAGAAAGGAGCAGAGGCGCAGAAAGAGTAGACAGGAGAAAGAAAGAAAGACAGAG 1028
Db 1165 GAGAAGAAAGGAAACAAGAGCCAGAAAGTAGACACAGAAAGAGCGGAAGAGCAGAG 1224
Qy 1029 CCAGCTCGAGGA-----GCAGGAGCCGGCAGAGAC 1061
Db 1225 GTTCCCTCCGAGAACTGACCGCTCCGAGCAAGCCCAACACAGAGCGCGCAGAAAGT 1284
Qy 1062 ACACACAGGCGCAGTTGTGACGACAGTACAGAGAGGTGGAGTGCCTTTTGGAAAGACAG 1121
Db 1285 GCCACAGCCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGGAGCAA 1344
Qy 1122 GTTGGTGACCTGGAGGATCGTCAGAGAGAGTGTGCTCCTTTGGCAACGGAAAGTGTTT 1181
Db 1345 GTCAGTGGCTCGCAGGAGACCTCTGAAAGAGAAACCTGCTCCGTTGGCGACAGAGTGTTT 1404
Qy 1182 GATGAGAAAGTGAAGCCCAAGAA-----GTTGTTGAGAGGTCAGCTGAGCAGCGTG 1238
Db 1405 GATGAGAAATAGAAAGTCCACCAAGAGAGGTTGTGGCCGAAGTCCAGCTCAGCAGCGTG 1464
Qy 1239 GAGAAGACAGAGGAGGAGCAGGAGGAGGAGAGGCTGAAGGGGGCGTGGTGTAGAA 1298
Db 1465 GAGGAGAGACCGGAAGAGCAGAAA-----ACGAGAGTGGAA 1500
Qy 1299 GGAAACAGGAGAAATCTTGGCCCTCGAGAAACTGGCTGAGCCCGCAGGAGGTCCCCAGGAA 1358
Db 1501 GAAACAGCAGGGTCTGTGCCAGCTGAAGATTGGTTGAATGGATGCAGAACCTCAGGAA 1560
Qy 1359 GCTGAGCCTGTGAGGAGCTGATGAAGACGAGAGAGATGTGTCTCTGAGGAGAGACAC 1418
Db 1561 GCCAAACCTGCAAGGAGCTGGTGAAGCTCAAGAAACGTCGTGTGTTCGGGAGAGGACCT 1620
Qy 1419 ACTCACTGACAGACCTAAGTCTGAGAGAGAGAGCTGCCCAACACCCAGAGGCAATT 1478
Db 1621 ACAAGGAGCTGACCTCAGTCTGATGAGAAAGGTGCTGTCNAAACCCCGAAGGCGTT 1680

Qy 1479 GTCAGTGAGGTGGAGATGCTGTCTCTCAGGAAAGAAATCAAGGTAACAGGAAAGTCCCTTG 1538
Db 1681 GTGAGTGGAGTGGAAATGCTGTCTCATCAGGAGAGAAATGAAGGTGCGAGGAAGTCCACTA 1740
Qy 1539 AAGAAACTCTTTCTAGTCTCAGGCTTAAAGAGCTGTCTGGGAAGAGCAGAGGGGAAA 1598
Db 1741 AAGAAGCTTTTACCGACTCTGGCTTAAAGAGCTTCTCGAAGAAACAGAAAGGAAA 1800
Qy 1599 CGAGGAGGTGGGGAGACGAGAGCCTTGGAGAAATACCAACACATTTCAACCGAATCCCCA 1658
Db 1801 AGAGG---GGAGGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATTTCTCCG 1857
Qy 1659 GAGAGTCTGATGAGCAAGAGGAGAGAGCTCTGCGTGTCTCCCGCAGAGGACCTTGAGAG 1718
Db 1858 GACAGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCTCTCATCCCTGAGGAGCCCGAGGAG 1917
Qy 1719 ACCACGTGTCTGGAGAAAGGGCGCTGGAAGCAACCCAGGATGGGAGCTGAGGAAGGA 1778
Db 1918 ATCAGTGTCTGGAAGAGGGCTTAGCCGAGGTGAGCAGGATGGGGAAGCTGAAGAAGGA 1977
Qy 1779 ACTACTTCCGATGGAGAGAAAGAGAGAGGAGTCACTCCCTGGGCTATCCTTCAAAAAG 1838
Db 1978 GCTACTTCCGATGGAGAGAAAGAGAGAGTGTCACTCCCTGGGCTTCAATTCANAAAG 2037
Qy 1839 ATGCTGACACCCAGAAACCGGGTCCGAAAGACCTTCTGAGAGTGACAAGAGGAGAAAGCTG 1898
Db 2038 ATGCTGACGCCCAAGAGCGTGTAGACGGCTTTCGGAAGTGTATAAGAAAGATGAGCTG 2097
Qy 1899 GAGAGGTCAAGAGCCACTTGTCTCTCACTGATAGCACAGTGTCCAGAAATCCAGAT 1958
Db 2098 GACAAGGTCAAGAGCGCTACTCTTGTCTTCCACCGAGAGCACAGCTCTGAAATCAAGAA 2157
Qy 1959 GAAGTCAAAACTGTGTGTGAGGAAACAAAAGCCAGAGAAACCAAGCGTAGGGTGGATCT 2018
Db 2158 GAATGAAGGAGCGTGTGAGAGCCAAAGCCGGAAGAACCAAGGCGCAAGTGGATACC 2217
Qy 2019 TCAGTGTCTTTGGAGACACTGATTTGTGTGCGATCATCCAAAGAGAGAGCAAGGAAGGCA 2078
Db 2218 TCAGTATCTTTGGGAAGCTTTAAATTTGTGTGGATCATCCAAAGAAAGAGCAAGGAGAGG 2277
Qy 2079 TCCTCTCAGATGATGAAGAGGGCCNAGGACACTGGAGGGGACAGTCAAGAGCAGAG 2138
Db 2278 TCCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGAGGAGACCAACCAAGAAAGTGT 2337
Qy 2139 GAGGCCAGCAAGACAAAGAACCGGAACAGACGCTGCTTCTGCCAGCACCCAGAGAGCAG 2198
Db 2338 GAGGCCGAAAAGACAAAGAGACGGGACAGACGGGATCCTTGTCTGTTTCCCAAGAACAT 2397
Qy 2199 GACCAAGCCGAAGGAAGTTCTCTCAACCGAGCCAGCGGGGAAGCCCTTCCGAAGGGGAAGGT 2258
Db 2398 GATCCAGGGCAGGGAAGTTCTCTCCCGAGCAAGCTGGAAGCCCTACCGAAGGGGAGGGC 2457
Qy 2259 GTCTCCACTTTGGGAGTCATTTAAAGATTAGTCACTCCAAAGAAAAAATCCAAGTCAAAA 2318
Db 2458 GTTTCCACTCTGGAGTCATTTAAAGGTTAGTCAAGCAAGAAAAAATCAAAAGTCCAAG 2517
Qy 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTGTCCACTGAG 2366
Db 2518 CTGGAAGAGAAAGCGAAGACTCCATAGCTGGGTCTGGTGTAGAACATTCCACTCCAGAC 2577
Qy 2367 ATCGAAACCGATGAGAGAAATCTTGGGTTTCCATTTAAGAAATTCATCCCGGACGGCGG 2426
Db 2578 ACTGAACCCCGTAAAGAAAGTCTCTGGGTCTCAATCAAGAAAGTTTATTTCTCGGACGAAG 2637
Qy 2427 AAGAAAGGCGACAGCGGAAGCAGAAACAAGCACTGTGGAAGACTCAGGCGCAGTGGAG 2486
Db 2638 AAGAAAGGCGCAGATGGGAAACAAAGAACAAAGCCCTGTTGAAGACGCGAGGCGCAACAGG 2697
Qy 2487 ATAAATGAGGACGACCTAATGTCCAGCGCTGTGCTCTGTCTGATGATTAATGCAAGT 2546
Db 2698 GCCAACGAAGATGACTCTGATGTCCCGCGCGTGGTCTCTGTCTGATGATGATGCTGTA 2757

OY	2547	GAGGGGAGAAAGATGC-----AAGCCAGGGGAAATACGACTCTCCACGTCTG	2557
Db	2758	GAAGGGAGAAATATGAGGAGCAGACAGAAAGCCAAAAAGGCGAGACAGCCGAGAGAAAG	2817
OY	2598	GGGGCTGTGATCGGTCTCCAGAGAGCTCAGTAAGACTCTGTGCCACACTGTGAGTTCGCA	2657
Db	2818	GCAGCCACTGAGTGTCTCAAAGAGCTCAGAGAGTCAAGTTCATATGATGGCAGACT	2877
OY	2658	GTCAATTGATGAGCACAGGGCAGTCAACCAATGTCTGAAGACGGGTCTCTTCTGTGATATCC	2717
Db	2878	GTCCGTGACGGGACGAGGGCAGCTACCATTAATGAAGAAAGGTCTCTTCTGTGATATCT	2937
OY	2718	GCTTCTCCTTAACAGAACTCTTGAACAACAGCGGGAGAAAGCCATGCCACTGTGAAGAG	2777
Db	2938	GCTTCAGTGAACAGAACTCTTGAACAAGTGAAGACTGAAGCCGCACTGTAACTGAAGAG	2997
OY	2778	GTCACTGAAAAAGCATCATTTGC---AGAAAGAACTCTGTCTCAACCGACGTTATCA	2834
Db	2998	GTATTGGAAAGAGAAAGTATTTGCAGAAAGAAAGACCCCAACGTTATCTGAACCTCTGCA	3057
OY	2835	GAGGGTAAAGATGCCCCATGTACGACATGTGTACCAAGTGAAGTGAATTTCACTCAGAACT	2894
Db	3058	GAGAAACAGAGAGGCCCGGGGCGACACGGTCTGTATGAGGCGGAAATGTACCCCGAAGCT	3117
OY	2895	GTGACAGCCACAGAGACTCAGAGGCTCTCCGATCTGAAGAAAGTTACCGAAGACTCGGG	2954
Db	3118	GTGACAGCTGTGAGAAACTGCAAGGGCCATTTGGGTGCGGAAAGAAACCGAAGCATCTCT	3177
OY	2955	GCCGAAGAGACCAAGACATGAGGTTCGCGAGTTTCCAGCTGACCTGACTCCCGACAGACC	3014
Db	3178	GCTGAGAGACCAAGAAATGTGTCAAGCAGTCTCCAGTTAAACGACTCCCGACAGACC	3237
OY	3015	ACAGAGGAAGCCACCCCAAGTTCAAGAGGTGAGAGTGTGTCTAGATACAGAAAGAG	3074
Db	3238	ACAGAGGAGGCACTCCGGGTGACAGAGGTGAAGGTGGGTACCTGACATAGAAAGAGCA	3297
OY	3075	GAGGCGCAGAGCGAGGCGCATCTTCAAGCCGTTGCGACAAGGTGAAGAGAGATGCCAG	3134
Db	3298	GAGAGGCGGCACTCAAGAGGTCTTCCACACAGTGGCAGAAAAAGTGAAGAGAAATCCGAG	3357
OY	3135	GTGCGCTGCAACCC---CAGACTGTGACAGAGAACGGGGTCAAAAAGCACTGAGAAAGTTGAG	3191
Db	3358	CTGCTTGCGCACCGGTGGGCGAAGAAATGTCTTCAAGCTGTGTGACAGAGACAGAGCGAGA	3417
OY	3192	GAGGTAGAGAGGACTCCCGAAGTGTCTGCTTCGAGAGAAAGAGAGCGTTATGCGGAA	3251
Db	3418	AGACCAAGAGAGCGAGCTGAAGCGTGGGTCTGAAAGAAAGACGGATGTAGTGTGAA	3477
OY	3252	GGACCCCTGACAGAACTGGAAGCTGAGACATCTTGGCACAGGGCTGTGACATGGAAGGCT	3311
Db	3478	GTAGATGTCTCAGAGGCGCAAAAAGCTGAACCTTTTATCACAGGGAAAGTGTGTGGGCGACAC	3537
OY	3312	ACTCCAGAGAGCTCTTGT---AAGTTCCTGAAGTCAACGCGAGATGTAGAC-----CAT	3359
Db	3538	ACCCAGAAAGCTTTGAAAAAGCTCTCTCAAGTCAACAGAGACCTAGAGTCCAGTGAAGTT	3597
OY	3360	GTGCGCACTGCGCA-----GGTTATCAAGCTCCAGCAGCTGATGAA	3401
Db	3598	GTAAACCAATTGTCAAGCGCAAACTTATGCTGGGGTAAATATCACAGAGAGATGTGATGAA	3657
OY	3402	CAGGCGCTGGCCCCCTGAGTCACTCCGAAAGCTTGAACAGACAGTGAACCAATGGAAGCACT	3461
Db	3658	CAGGCTATCCCCCTGACTCGTGTGAAGAACCCCTTACAGACAGTGAAGTCACTATGGAAGCAC	3717
OY	3462	CCCTTAGACAGTTTCAAGACCTGCAAGTGAAGGGAACAACAAGATGAAACATTTGACAGCAG	3521
Db	3718	CCCGTAGCCGACTTTGACGACCAAGGCACAACCCAGAAAGACAGATTGTGAAATTCAT	3777
OY	3522	GACAGTAAAGCACTGCAAGCTGTCAAGGCACTCAAGGTCAACAGAAAGAAAGCGCGCTACT	3581
Db	3778	GAGGAGAAATGAGTCTGCATCTGTGATCCCACTACAGGGGGCAGAGAAACAGACGCAATTTCT	3837
OY	3582	GCTCAGAAAGAGAGCTTTCGACATTAATATATGTTTCAGCCCGCAGAGAACATATGGG	3641

Db	3838	GCACAGAAAGAGGCGCTCCAGACCTTTCAGTTTGTGTTCCAGGAAGAACTTAAAGAA	3897
Qy	3642	GAAGAACAGAGAAAGAGTGTCTTGAACCTACACAGCAAGAGCTTACTGTCAGCCGTG	3701
Db	3898	CAATCAAAAGATGAGAGACCTCAGAGCATACAGTAAAGAGGTGTCAGTGAAGAACTGTA	3957
Qy	3702	CCCGTTTGGCCAAAGACTAGGTGGGTCAAGGGGTGA-----GGTTGACTGGTTG	3752
Db	3958	TCCATTCTGTCAAAAGACTAGGGGACTCAAGAGGCTGACAGTATGCTGATGAAGAAACC	4017
Qy	3753	GATGAGAAAGATGCTCAAGAGAAAGACAGAGGAGTGTGATACCTCG-----ACCCAAC	3806
Db	4018	AAAGACGTACATTTTTCGAAAGACTTGAAGGGGTCTATAGACACAGGCATTAACGTAGT	4077
Qy	3807	AGTCMAAGGCTCTGATGTGACATATGACAGTGAAGTATGGAGTGGCCGGGTGCAG	3866
Db	4078	CGGAAAGAGTCACTGAAGTTGCCCTTAAAGGTGAAGGAGCAGAAAGAGCTGAATGTAAA	4137
Qy	3867	GAAGAAGAGATCTAGTAGTCAGAG-----TCTTAGCTGTGAGAGAGGA	3911
Db	4138	AAGGATATGCTCTTGAACCTGACAGATCAAGTAACTCTCCATCCCGCTGTGAAGA	4197
Qy	3912	GAGATGGAATCTGACGTTGAAAAGAGAAAGAGAGCAAAAGCCAGACAAAGTGAATGA	3971
Db	4198	GAGATGTGATGTTCAAGTCGAAAGGAGAAACAGAACAGAGCCAACTGATGAATGAA	4257
Qy	3972	GAAAG---TGAGCAGAGAAACAGCCGCTCTCTGAGCATGAAGAACCTACGGAAACCAGTC	4028
Db	4258	GAGAAAGCTTAGCAGCAAGAAACAGCTGTTTACGATCTGAAGAGGTCAATGAACAGCTCCTC	4317
Qy	4029	CTGACACTTGAATGCCCACTCAGAGAGGGGAAAGGCATCGGAAAGCTTTGAGAGAAC	4088
Db	4318	CAGACAGTGAATGTGCCATCTATGATGGGGCAAGAAAGTCAGCAGTTTGGAGGAAGC	4377
Qy	4089	CTTCTC---TCCACAGCAAGCAAAAGCAGGTTCATAGAGGTTTCAAGTTCAAAAGCTG	4145
Db	4378	CTCTCTCCCTGCTAGGTCAAGAGAGGAGCAGTATACCAAAATTCAGATTCAGAGCTCT	4437
Qy	4146	GACACAAACAGTCACTCAAAACAGCAGAAAGCTTGAAAA-----GGTCATGAAGACGGTT	4199
Db	4438	GAGGCACTCATCTCACTCAACAGGCGCTGCAGAGAGGAAAGGTCTTAGAGAAAACTGCC	4497
Qy	4200	GTAATTTCAAGACAGGTGAAGTCCAGATGTGTAGGTGCACACTTATTCACAGCTGAG	4259
Db	4498	AACATTTTAAAGAACAGGTGAAGCTTGGAGCCTGCAGGTGCACATTTAGTTCTTGAAGAG	4557
Qy	4260	AAGTCTCTGCAACGGGTGGCCACTGGACTTTCAGCAATGCAGAGGACAGCGTACCCCTG	4319
Db	4558	AAATCTCTTAAAAAAATGAAGACTTTGGCGCTCATTCAGGGGAAGATGCTGTGCCACA	4617
Qy	4320	GGGCGTAGTCTCAGGCAGATTCATCCCAATCATATGTAATCTCTGCTCTGAAGACAC	4379
Db	4618	GGGCGCAGCTGTCAAGGCAAAATGACACACAGATATGATGTGCTACTACCAAGAAAGGC	4677
Qy	4380	CTACATCTCTGACTCAAGAGAAATTAAGCGATCCCAAGAGAGCGATCAGAGAAAG	4439
Db	4678	TTTAGTTCCACCTGGAAGAGAGAAACCAATCATCTGAAGTGAAGTCCAGATGAAGTCC	4737
Qy	4440	GACAAGCAGATGTGTGCTGTATGCTGACGGCAGAGAGAGTACAGCAATGAAAAAGTCC	4499
Db	4738	GATGAGAGGTTGCTTGGC-----AGGAGGTCAAGTGAAGTGAAGTGAAGAGAT	4791
Qy	4500	CTCAAGCTGAACCTGAGATCTGGAATCTGGAAGTGAAGAGCAACAAGTTGTCTGAAC	4559
Db	4792	TTAAGCTTAAAAATGGATTTTGGAACTTGGACCAAAAGACAGTAACTGTCCAAAC	4851
Qy	4560	GTCATTGACAGAGCGTTGACCAATTCCAGCTAC---AGAAACAGCCCCGAAACTCAT	4616
Db	4852	ATCATTCAGACAGCGCTGTGACAGTTTGAAGTACAGAAAGAAACAGCCACGAATGTGG	4911
Qy	4617	GCTTATGATTCACAGACCCAGGTTCT-----CTGATCCAGGCTTGACAG	4660

Db 4912 ACGTCTGAGTTACAGACACAAGCTCACGTGATATAAAGCTGACAGCCAGGAGCGCTGGACAG 4971
Qy 4661 CAGGAGGCCAAACAGATGCTGGACAAATAAGATGCGCAAGATGAAACACCCAGTGCC 4720
Db 4972 GAAACGGAGAAAGAGAGAGAACTCAGGCCCTCTGCACAGGATGAAACACCAATTA 5031
Qy 4721 GCAGCCCA-----GAGAGACTTGCAGTCTCTGACCGTCTCGGA 4759
Db 5032 TCAGCCAAAGAGAGTCAAGTCAACCGCAGTGGGCAAGCACATCTCTGATATTTCAAA 5091
Qy 4760 GGCAATGGGCTCAGCCTCGGAAATGCTTCCCGCTTGCAGTTGAAAGCGCGGTGTCAA- 4818
Db 5092 GACATGAGTGAAGCCTCAGAAAGACCATGACTGTTGAGGTAGAAGTTCACATGTAAT 5151
Qy 4819 ----- 4818
Db 5152 GATCAGCAGCTGGAAGAGGTCGTCTCCATCTCAGGAAGAGGAGGTGGAGCTGGAACA 5211
Qy 4819 -----AGTAAGCATTGAGAACTGCCT 4840
Db 5212 AAGTCTGTCCAGAAAGATGATGGTCATGCCCTTGTAGCAGAAAGAAATAGAGAACTCA 5271
Qy 4841 CCTCAACCCAAAGA--TCCAAAGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTA 4898
Db 5272 GTTGAACCGAAAGAGATGAAGAGGTGATGATGATGACCTTGAAACCCAGAACTCA 5331
Qy 4899 GCCCAGCAGAGGCCAGTGCCTCTGGAACCTTAACAAAGAAATCCCGACAGACCAACGGA 4958
Db 5332 GCCCTGGCTGATCTGATGCCCTCAGGAGGCTTAACCAAGAGTCCCGCAGATACAAATGGA 5391
Qy 4959 CCAAGCTAACCGAGGAGGGGATCCCGCAAAAGTTGAGGTCCAGGAAGAAGAAATGTCT 5018
Db 5392 CCAAAACAAAAGAGAGAGGAGTGGCCAGGAAGTAGAAATGTCAGGAAGGAAAGTGCAC 5451
Qy 5019 ACCAAGTCAGTCAAGAGAGAACAAGGCCAGCAGACAGAGACCTGCAGGAGCCAAAGGGA 5078
Db 5452 AGTGAATCAGATAAAGCATCACCCCAAGCAGAGAGGAGTTACAGAAACAAGAGAGA 5511
Qy 5079 GACCTGCAGAAATCCTAAGATGTTAGTTG-----CTCATTTGTAC 5117
Db 5512 GAATCTGCAAGTCAGAACTTACAGAACTTAAACATCATGCAGATTAAACTCATTTGCT 5571
Qy 5118 ATCTGTAAGACCAAGATGTGAACCAAGTACAGAAACA-----GATGCTGCTGGGAC 5173
Db 5572 GTTTGGAAGACCAAGATGTGAAGACAAGTAGTAGAAGAAATGAATGCTGCTGAGAG- 5630
Qy 5174 CTTGAGACCAAGATTTTCAGAGCCCATCAGATCCAGAGAGAGGCGCTGCCAATGATTTCC 5233
Db 5631 CTGAGACCAAGTATTTTCAAG-ACTTTGAGAAATTGGAGAGAGGACATCACTGATCTCA 5689
Qy 5234 ACCAGTAGACACCCGACAAATTTCTGAGGCTTCATCGGAGCTTAGAGCCAGCTAAACATT 5293
Db 5690 TTTCTAGAGAGC-CCCTGACAACTCTGAGGCTTCATCAGAGCTAGAGCCATTAAACATT 5748
Qy 5294 TCCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGTAGTGCCTGCGTGTATTTTCAACATT 5353
Db 5749 TCCTCTTTTCAAGACCAACCTACAATTTTCCCTTGTAAAC----- 5788
Qy 5354 TCCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGTATGTCGCTGATTTTCGATTTA 5413
Db 5789 -----CATATAAATTTCTGATTTA 5806
Qy 5414 AGGTCTCGCTTTCACACCTGGAAC-CAATTCGCCATACCTAGTTCACCTTCTCAAACT 5472
Db 5807 AGGTCTTAATTTCTTAACCTGGAAGTGGAGTTGGCAATACCTAGTTCGCTTCTGAACCT 5866
Qy 5473 GGAGCATCCTCTTTATGATATATGATGTTTTATGTTTATGTTAGTCCCTCCTCGTACCTAT 5532
Db 5867 GGAGTATCACTCTTTTACATATTTATATGATGTTTTTAAGTAG---TCTCCTGCTATCTAT 5923
Qy 5533 TGATATATTTTCTTAAGCTTT-----AAGCATGCTTTTGTATTA 5575
Db 5924 TGATATATTTTCTTAAGCTTTAAGAAATGTGCAGGATACATGCTTTTTTGTATCA 5983

Qy 5576 TGCAATATATAACGGGTGTGCAGCCATAGCAGCGCTTTTGAAGCTCCAAGCCTCAACTG 5635
Db 5984 CACAGTATATGATGGGCGATGTGCCATAGTGCAGGCTTGGGAGCTTTAAGACCTCAGTTA 6043
Qy 5636 TAACCTGCAGCAACAGATAAACATTC-----CTGGCAAGAGAGACAGTCTTTTT 5686
Db 6044 TATAACCCACAAAACACAGAGCCTCTAGATGTAAACATCTCTGATCAAGGTACAAATCTT 6103
Qy 5687 TAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCCCTCTGAAAGTGGTTGTTTCTCT 5746
Db 6104 TAAATTCATTAATGATTTGAGGTCCATATTTAGTGGTACTCTGAAATTTGTCATCTTCT 6163
Qy 5747 ATGCACAGCAGCTCAGAAATAAAACCCCATTTTGAACATCCAGATGTCCTCAATATT 5806
Db 6164 ATTACACGGAGTGTGCCAAAACATAAAAGCATTTTGAACATACACAGAAATGTTCTATTGTC 6223
Qy 5807 ACCATGATTTTTTCCCCCTTTTGTCTTAATCCAGTCCAGGTGGAAAGAGTCTCTCTG 5866
Db 6224 ATTGGGAAATTTTGC-----TTTCTAACCCAGTGGAGGTAGAAAGAAAGTTATATTCT 6276
Qy 5867 TGTGAGA-----TTAAGCCCTGTCTCTTTAAATGATATGACAAATCAGT 5909
Db 6277 GTTAGCAATTAACCTTTTACATCCTTTTCTACTTGTATGTTGTTGGACCGATTAAGT 6336
Qy 5910 GTGCTTAAGCCCATGAGA-----TGTTCTTAATGCAGAAAGAAATCTGT 5953
Db 6337 GTGCTTAATCCTGAGGCAAGTAGTAATATGTTTATATGTTATGAAGAAAGAAATGT 6396
Qy 5954 TGTACGTTTTTTTGTATGTTACTCTTATCTGAGACCGAAATTCATATGCAGATGCAAGTG 6013
Db 6397 TGTAAAG--TTTTGATTTACTCTTATATGCTGAGCTGCATTCACACATGCGCATGAAATA 6454
Qy 6014 AGTCCTGTTCTTTACAGATGTTATTTGATAGACTGGAAGTTTGTCTGTGTTATATCTG 6073
Db 6455 AGTCAGTTCTTTTACAAATGTTATTTGATAGATCTGGAATGTTGTTGTCATATTG 6514
Qy 6074 TGCCCTTCTTTTAAAGCAATGTTGATTTATGTTTCTTTGGATAAAATTTGATTTGACAA 6133
Db 6515 TGCCATTTCTTTTAAAGCAATGTTGCAACATTTCAATTTGATAAGTTGTTGATTGACGA 6574
Qy 6134 CTGATTTAAATAAA 6147
Db 6575 CTGATTTAAATAAA 6588

RESULT 14

US-60-245-228-517/c

; Sequence 517, Application US/60245228

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CLO00878

; CURRENT FILING DATE: 2000-11-03

; NUMBER OF SEQ ID NOS: 630

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 517

; LENGTH: 6607

; TYPE: DNA

; ORGANISM: HUMAN

US-60-245-228-517

Query Match 35.4%; Score 2178; DB 68; Length 6607;

Best Local Similarity 64.4%; Pred. No. 0;

Matches 4222; Conservative 0; Mismatches 1795; Indels 537; Gaps 43;

Qy 3 GGCGGGGGGAGTAGAGACCACTGAGCCATGGCGCAGCAGTTCACCGAGCAGCG 62

Db 6443 GGCTAGCGCGGAGAGTGCAGGAGCCATGGCGCGCGGAGCTCCACCGAGCAGCG 6384

Db	5321	GCCACGAGCCCCGGTTATCAGCTGAATATGAGAAATTGACTGCTCCCTTGACAGAGACCA	5262
Qy	1122	GTTCGTACCTTGAGGCGATCGTCAGAGGAAGTGTGCTCTTTGGCAACGAAAGTGT	1181
Db	5261	GTCAGTGGCTCGCAGGGAACCTTCTGAAAGAAACCTGCTCCGTTGGCAGCAGAAAGTGT	5202
Qy	1182	GATGAGAAATGAGACCCCAACCAAA--GTTGTTGCAGAGTCTCAAGTGAACCGTG	1238
Db	5201	GATTAGAAATTAAGATGCCACCAAGAAAGATTGTGGCCGAAGTCAAGTCAAGACCGTG	5142
Qy	1239	GAGAAACAAGAGGAGCAGGGAGGAGAGAGAGGCTGAAGGGGGGCGTGTGGTGA	1298
Db	5141	GAGGAGGAACCGAAGAGCAAAA-----ACGAGGTGGA	5106
Qy	1299	GGAACAGAAATCTTGGCCCCCTGAGAACTGGCTGAGCCCAAGAGGTCCCCAGAA	1358
Db	5105	GAACACACAGGCTGTGGCAGCTGAAGAAATTGGTTGAATGGATGCAAACTCAGAA	5046
Qy	1359	GCTGAGCTTGCTGAAGACTGATGAAGACAGAGATGTGTCTTGGAGAGACAC	1418
Db	5045	GCCGAACTCCCAAGAGAGCTGTGAAGCTCAAGAAACGTGTGTTCCGGAGAGACCT	4986
Qy	1419	ACTGAATGACACACTAAGTCTCTGAAGAAAGCGTCCCAACCCAGAAAGGCACT	1478
Db	4985	ACACAGGAGCTGACCTCAGTCTGTATGAAAGGTGTCTCAAAACCCCGAAGGCTT	4926
Qy	1479	GTCAGTAGGTGAGATGTGCTCTCTCAGAGAAATCAAGTCAAGGAAAGTCCCTTG	1538
Db	4925	GTGAGTAGGGTGAATCGTGATCAGAGAGAAATGAAGTCAAGGAAGTCACTA	4866
Qy	1539	AAGAACTCTTCACTAGCTCAGGCTTTAAAGAGCTGTCTGGAGAGAGAGAGGAAA	1588
Db	4865	AAGAACTTTTTCACAGCACTGCTTAAAAAGCTTTCGAAAGAAAACAGAAAGGAAA	4806
Qy	1599	CGAAGAGTGGGGGAGACGAAAGCCTGGAGAAATCCAACATTCACACCGAATCCCCA	1658
Db	4805	AGAAGA--GAAGAGACGAGGAATAGGGAGACACACTCAAGTTCAGCGATTCTCG	4798
Qy	1659	GAGAGTCTGATGACGAAGGAGAGAGAGCTGTGCGTGTCTCCCGAGAGAGCTTAGAG	1718
Db	4748	GACAGCAGAGAGAGCAAAAGGGCGAGAGCTGTGCTCATCCCTGAGAGAGCCGAGAG	4689
Qy	1719	ACCACTGTCTGAGAAAGGGCCGCTGAGACACCCAGAGATGGGAAAGCTGAGAGGA	1778
Db	4688	ATCCAGTGTCTGGAAGAGGCTTAGAGCGATGACAGAGATGGGAAAGCTGAAAGAGA	4629
Qy	1779	ACTACTTCGATGAGAGAAAGAGAGAGAGGATACCTCCCTGGGCAATCCTCAAAAG	1838
Db	4628	GCTACTTCGATGAGAGAAAGAGAGAGAGGATCTCCTGGGCAATCATTCAAAAG	4568
Qy	1839	ATGTGTACACCCAAAGACGGGTCCGAACCTTCTGAGATGACAGAGAGAGACCTG	1898
Db	4568	ATGTGTACGCCAAGAGAGCTGTTAAGCGGCTTGGAGAAAGATGAAGAGATGAGCTG	4509
Qy	1899	GAGAAAGTCAAGACGCCACCTTGTCTCTCACTGATAGCAACGTGTCAAAATGCAAGT	1958
Db	4508	GACAAAGTCAAGAGCGCTACCTTGTCTTCCACCGAAGACACGCTTGAAATGCAAGA	4449
Qy	1959	GAAATCAAACTGTGTGTGAGGAACAAAGCCAGAGAACCAAGCGTATGGGTGATCT	2018
Db	4448	GAAATGAAGGAGACGTGGAAGACCAAGCCGGAAGAACCAAGGCGAAGGTGATACC	4389
Qy	2019	TCAGTGTCTTGGGAAGCATGATTTGTGTGATTCATCCAGAAAGAGACAGGAAGCA	2078
Db	4388	TCAGATCTTGGGAAGCTTTAATTTGTGTGATTCATCAAGAAAGAGCAAGAGAGG	4322
Qy	2079	TCTCTTCAGATGATGAAGAGGAGCCAGAGCACTGGAGGGGACAGTCAAGACAGAG	2138
Db	4328	TCTCTTCATGATGAGAAAGGGGAGCCAAAGCATGTGGAGAGAGCACACAGAAACTGAT	4265
Qy	2139	GAGGCAAGAAAGCAAGAAAGCCGGAACAGAGCTGTCTGCGACAGCCCAAGAGAGAG	2199

Db 4268 GAGCGCGAAGAACAAAGAGACGGGGACAGACGGGATCCTTGTGGTTCCCAAGAACAT 4209
Qy 2199 GACCAGCGCAAGGAAGTTCTCACCAGCAGCGGGAAGCCCTTCGGAAGGGGAAGGT 2258
Db 4208 GATCCAGGCGCAGGGAAGTTCTCCCGGAGCAAGCTGGAAGCCCTTACCGAAGGGAGGCG 4149
Qy 2259 GTCTCCACTTCGGAGTCATTTAAAGATTAGTCACTCCAAGAAAAAATCCAAGTCAAAA 2318
Db 4148 GTTTCACCTGGGAGTCATTTAAAGGTTAGTCAGCCCAAGAAAAAATCAAGGTCCAAG 4089
Qy 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGACAGTGTGTCACCTGAG 2366
Db 4088 CTGGAAGAGAAAGCGAAGACTCCATAGCTGGGCTGTGTGTAGAACATTCACCTCCAGAC 4029
Qy 2367 ATCGAAACCGAGTAGAAGAAATCTTTGGGTTTCCATTAAAGAAATTCATCCCGGACGGCG 2426
Db 4028 ACTGAACCCGGTAAAGAAAGATCTCTGGGTCTCAATCAAGAAAGTTTATTCCTGGACGAAGG 3959
Qy 2427 AAGAAAAGGCGAGCGGAGCAGCAAGCAAGCCACTGTGGAAGACTCAGGSCCAGTGGAG 2486
Db 3968 AAGAAAAGGCGAGATGGGAAACAAAGAACCAAGCCCTGTTGAAGACGCGAGGSCCAACAGGG 3909
Qy 2487 ATAAATGAGGACGACCCCTTAATGTGCCAGCGTGTGTCCTGTCTGTAGTATAATGCACTG 2546
Db 3908 GCCAACGAAGATGACTCTGATGTCCCGCGCTGTGTCCTCTGTCTGAGTATGATGCTGTA 3849
Qy 2547 GAGAGGGAGAAAGATGG-----AAGCCCAAGGGGAATACGGAGCTGCCAGCTGTG 2597
Db 3848 GAAAGGGAGAAAAATGGAGGCGACAGCAAGSCCAAAAAAGCGCAGAGCAGCCGAGCAGAG 3789
Qy 2598 GGGCTGTGTACGTGTCCGAGGAGTCACTAAGACTCTGTGTCACACTGTGAGTGTGCGCA 2657
Db 3788 GCAGCCACTGAGGTGTCCAGGAGCTCAGCGAGAGTCAGGTTTCATATGATGGCAGCAGCT 3729
Qy 2658 GTCAATTGATGGGACAGGCGAGTCACCACTGTGCAAGACCGTCTCTCTGTGCGATATCC 2717
Db 3728 GTCCTGACGGGACAGGGGAGCTACCAATTAATGAAGAAAGGTCTCTCTCTGTGATATCT 3669
Qy 2718 GCTTCGGTAAACAGAACCTCTTTGAACACACACAGCGGGGAAAGCCATGCCACCTGTTGAAGAG 2777
Db 3668 GCTTCAGTGACAGACCTCTTTGAACAAAGTAGAAGCTGAAGCCGACACTGTTAACTGAGGAG 3609
Qy 2778 GTCACTGAAAGACATCATTTGC---AGAAGAACTCTGTGCTCACCCAGACGTTTACCA 2834
Db 3608 GTATTGAAGAGAGAAATTTGCAAGAGAAAGAACCCCCACCGTTACTGAACTCTTGCCA 3549
Qy 2835 GAGGGTAAAGATGCCATGACGACATGTCACCACTGCAAGTGAAGTGAATTTACCTCAGAGCT 2894
Db 3548 GAGAACAGAGGCGCGGGGCGACACGGTCTGTAGTGAAGCGGAATTGACCCCCGAAGCT 3489
Qy 2895 GTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAAAGTTACCGAAGCATCGGG 2954
Db 3488 GTGACAGCTGCAGAAACTGCAGGGCCATTGGGTGCCGAAGGAACCGAAGCATCTGCT 3429
Qy 2955 GCCAAGAGACCAAGACATGTTGTCGAGTGTCCAGTGTTCACAGCTGACTGTCTCCCAAGAC 3014
Db 3428 GCTCAAGAGACCAAGACAAATGTTGTGACGAGTCTCCAGGTTAACCGACTCCCAAGACACC 3369
Qy 3015 ACAGAGGAAGCCACCCAGTTTCAGAGGTAGAGAGTGTGTGTAGATACAGAGAAGAG 3074
Db 3368 ACAGAGGAGGCCATCCGGTGCAGAGGTGAAGGTGGCGTACTGTGACATAGAAGACAA 3309
Qy 3075 GAGCGCCAGACGCGAGGCCATCCTCAAGCCGTTTCAGACAAGGTGAAGAGGAGTCCAG 3134
Db 3308 GAGAGCGGACTCAAGAGGTCTCTCAGGCACTGCAGAAAGTGAAGAGGAATCCAG 3249
Qy 3135 GTGCTCGAACCC---CAGACTGTGCAGAGAACCGGGTCAAAAGCACTGTGGAGAAGTTGAG 3191
Db 3248 CTGCTGGCACCGTGGGCCCAAGAGATGTGCTTCAGCCTGTGACAGAGCAGAGGCAGAA 3189
Qy 3192 GAGTAGAGGAGGACTCCGAGTCTGGCTTCGAGAAAGAGAGGAGCTTATGCCAAA 3251
Db 3188 AGACCAAGAGAGAGGCTGAAGCGTCGGGTCTGAAGAAAGAGAGCGGATGTAGTGTGAAA 3129

Qy 3252 GGACCCGTGCAGGAAGCTGGAGCTGAGCATCTTGCACAGGGCTCTGAGACTGGACAGGCT 3311
Db 3128 GTAGATCTCTCAGGAGGCAAAAACTGAGCCCTTTTACACAAGGAAGGTGGTGGGCGACAGCC 3069
Qy 3312 ACTCCAGAGAGCCCTTG---AAGTTCTCTGAAGTCAAGGAGATGTAGAC-----CAT 3359
Db 3068 ACCCCAGAAAAGCTTTGAAAAAGCTCTCAAGTCAAGAGAGCATAGAGTCAGGTGAGCTT 3009
Qy 3360 GTCGCCACGTGCCA-----GGTTATCAAGCTCCAGCAGCTCATGAA 3401
Db 3008 GTAACCACTTGTCAAGCCGAAACCTTAGCTGGGTAATAATCAAGGAGATGGTATGAA 2949
Qy 3402 CAGCCCTGGCCCCCTGAGTCAATCCGAAACCTTGCACAGACTGAGACAAATGGAAGCACT 3461
Db 2948 CAGGCTATCCCCCTGACTCGGTGGAACCCCTACACAGACTGAGACTGATGGAAGCACC 2889
Qy 3462 CCCTTACAGATTCAGACACTGACGATGGGACACAGCAAGATGAACCAATGACAGCAG 3521
Db 2888 CCGTAGCCGACTTGAACGACCCAGGCAACCCAGAAAGACGAGATTGTGGAATCCAT 2829
Qy 3522 GACAGTAAAGCACTGTCAGCTGTCCAGCAGTCAAGCTCACAGAAAGAGCGGCTACT 3581
Db 2828 GAGGAGAAATGAGTCCGATCTGGTACCCAGTCAAGGGGACAGAAAGCAGAGGCACTTCT 2769
Qy 3582 GCTCAGAAAGAGGAGCCTTCGACACTACCTAATAATGTTCCAGCCCAAGGAACATGGG 3641
Db 2768 GCACAGAAAGAGAGGCTCCAGCACCTTCCAGTTTGTGTTCCAGGAAGAACTAAAGAA 2709
Qy 3642 GAAGAACCAAGAGAGATGTTCTTGAACCTTACACAGCAAGACTTACTGTGACGCGTG 3701
Db 2708 CAATCAAGATGGAAGACACTTAGAGCATACAGATAAGAGGTGTCAAGTGAAGTAA 2649
Qy 3702 CCGTTCCTGCAAGACTGAGGTGGGTCAAGAGGTGA-----GCTTACTGCTG 3752
Db 2648 TCCATTCTGTCAAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCTGATGAAGAAACC 2589
Qy 3753 GATGGAGAAAAAGTCAAAAGAAAGCAGGAGTGTGTTGTACACTCTGG-----ACCCAAC 3806
Db 2588 AAAGACGTACCAATTTTCGAAGACTTGAAGGCTCTATAGACACAGGCATACAGTCACT 2529
Qy 3807 AGTCAAAAGCTGCTGATGTGACATATGACAGTGAAGTGTGGAGTGGCGGCTGTCAG 3866
Db 2528 CGGGAAGAGGTCACTGAAGTTGCCCTTAAAGGTGAAGGGACAGAAAGCTGAATGTAAA 2469
Qy 3867 GAAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCTGGAGGAGGA 3911
Db 2468 AAGGATGATGCTCTTGAACCTGCAGAGTCAAGTAAAGTCTCTCCATCCCGCTGGAGAGA 2409
Qy 3912 GAGATGAAACTGACGTTGAAAAAGGAGAAAGGAGACAAAGCCAGACAGCAAGTGAAGTAA 3971
Db 2408 GAGATGATGTTCAAGTTCGAAAGGGAGAAACAGAGCAGAGCCACCCATGTGAATGAA 2349
Qy 3972 GAAGG---TGAGCAGGAACACAGCCGCTCTGAGCATGAAGGAACCTACGGGAAGCCAGTC 4028
Db 2348 GAGAAGCTTGAGCACCAAGACAGCTGTACCGTATCTGAAGAGGTCAAGTAAAGCAGTCTCTC 2289
Qy 4029 CTGACACTTGACATGCCAGCTCAGAGGGGGAAGGCACCTGGAAGCCTTGGAGGAGC 4088
Db 2288 CAGACAGTGAATGTGCCCATCATAGATGGGGCAAGGAAGTCAAGCAGTGTGGAAGGAGC 2229
Qy 4089 CTTTCTC---TCCAGACCAAGACAAAGCAGGTTGATAGAGGTTCAAGTTCAAAGCCTG 4145
Db 2228 CTTCTCTCCCTCCTAGGTCAAGAGGAGGCAAGTATGCCAAATTCAGTTCAGACTCT 2169
Qy 4146 GACACAAAGTCACTCAAAACAGCAGAAAGCTGTGGAATA-----GGTCATAGAAAACGGTT 4199
Db 2168 GAGGCATCATTTCACTCTAACAGCGGCTGCAGAGGAGGAAAAGGTCTTAGGAGAACTGCC 2109
Qy 4200 GTGATTTTCAGAGACAGGTGAAGTCCAGAGTGTGTAGGTGCAACATTTATCCAGCTGAG 4259
Db 2108 AACATTTTGAAGAACAGGTGAAGCGTTGGAGCCTGCAAGGTGCACATTTAGTTCTTGGGAAG 2049

```

QY 4260 AAGTCCTGCAAGGGTGGCCACTGACTCTTCAGCATGCAAGAGACACGGTACCCCTG 4319
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2048 AAATCTCTGAAAAAAGAAAGACTTTGCCGCTCATCCAGGGGAAGATGCTGTGCCACA 1989
QY 4320 GGGCCTGAGTCACAGGAGCAATCCATCCCAATCATATGTAATCTCTGCTCGAAGACCC 4379
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1988 GGGCCCACTGTACGGCAAAATCCAGCACAGTATGATGTCGTACTACCAAGAAAGGC 1929
QY 4380 CTACATCTCTGACCTTACAGAGAAATTAAGCGCATCCAGAGAGACGATCAGAGGAAG 4439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1928 TTAACTTCGACCTGGAGAGAGAAAAACACATCATGAAGTGAAGTCAGATGAATGTC 1869
QY 4440 GACAAAGCCAGATGCTGTGTCCTGATCTGACGCAAGAGAGATGACGCAATGAAAAATGC 4499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1868 GATGAGCAGGTGTGCTTGCC-----AGGAGGTCAAGAGTGTGCAATGTAGAGAGAT 1815
QY 4500 CTCAGAGGTGAACCTGAGATCCTGGAACCTTGAGAGTAAAGCAACAAGTTGTGTCGAAC 4559
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1814 TTAGAGCTTGAATAATGGATTTTGGAACTTGAGACCAAAAGAGTAACTTGTCCAAAAC 1755
QY 4560 GTCAATTCAGACAGCCGTTGACCACTGCGACCGTAC--AGAAACAGCCCCCGAAACTCAT 4616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1754 ATCATCCAGACAGCCGTTGACCACTTGTACGTACAGAAAGAACGCCAACGAAATGTTG 1695
QY 4617 GCTTATGATTCACAGACCCAGGTTCT-----CTGCATGCAAGCTTGAACAG 4660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1694 ACGTGTGAGTTACAGACACAAAGCTCAGTATMAAGCTGACAGCGACAGACCTGACAG 1635
QY 4661 CAGGAGGCCCAACAGATGCTGCAAAAATGAAAGATGCGAAGATGAACACCCGATGCC 4720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1634 GAACCGAGAAAGAGAGAGAACCTCAGCCCTCGCACAGAGATGAACACCAATTAAT 1575
QY 4721 GCAGCCCA-----GAGAGACTTGAAGTCTGACCGTTCTGGA 4759
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1574 TCAGCCAAAGAGAGATGAGATCAACCCGACGTGGGCAAGACATTTCTGATATTTCCAAA 1515
QY 4760 GGCATGGGCTCAGCTTCGAAATGCTTGCCTGCGCTTGCAGTTGAAAGCCCGGTGCA- 4818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1514 GACATGAGTGAAGCTTCAGAAAAAGACATGACTGTTGAGGTGAAGTTCCAATTAAT 1455
QY 4819 ----- 4818
Db 1454 GATCAGCAGCTGGAAGAGTCTCTCCCATCTGAGAGAAAGAGAGTGGAGCTGGAACA 1395
QY 4819 -----AGTAAGCAATTGAGAGCTGCT 4840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1394 AAGTGTGTCCAGAGATGATGTGATGCTTGTAGAGAAAGATTAAGAAAGTCACTA 1335
QY 4841 CCTCAACCCAAAGA--TCCAAAAGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTA 4898
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1334 GTTGAACCGAAAGAGATGAAAAAGGTGATGATGTTGATGACCTGAAAAACCAAGACTCA 1275
QY 4899 GCGCCAGGAGAGCCAGTGGCTCTGGAACCTAACCAAGAAATCCCCAGACCAACGGA 4958
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1274 GCGCTGGCTGATATGATGCTCTCAGAGGCTTAACTCAAAAGAGTCCCAATACAAAATGGA 1215
QY 4959 CCAAGCTTAAACGAGAGGCGCATCCGCCAAAAGTTGAGGTCCAGAGAAAGAAATGTCT 5018
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1214 CCAAAACAAAAGAGAGAGAGATGCCAGAAAGTGAATTTGAGAGAAAGAAAGTGAC 1155
QY 5019 ACCAAGTCAGTCAAAAGAGAACAAAGCCAGGCAAGAGAGACCTGCAAGAGCCAAAGGGA 5078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1154 AGTGATCAGATTAAGCGATCAACCCCAAGCAAGAGAGAGTTCAGAAAACAAAGAGAGA 1095
QY 5079 GACCTGGAGATCTTAAGATGTTAGTTG-----CTCATTTGATC 5117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1094 GAATCTGCAAGTCAGAACTTACAGATCTTAAACATCATGCAAGTAAATCATTTGTCT 1035
QY 5118 ATCTGTAAAGCAAGATGTGAACAAAGTCAAGAACAA-----GATGCTGTGTTGGAGC 5173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1034 GTTTGGAAAGCAAGATGTGAAGCAAGTAGTAAGAAAGAAATGAAATGCTGTGCTGAGA- 976
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5174 CTTGAGACCAAGATTTCAAGGCCATGAGATCCAGAGAGAGGCGCTCCAAATGATTTCC 5233
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 975 CTGAAGACCAATTTTATAGA-ACTTTGAAGATTTGAGAGAGAGAGACATCAATGATCTCA 917
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5334 ACCAGTAGAGACACCCGACAAATTTGAGGCTTCATCGGAGCTAGAGCCAGTAACTT 5293
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 916 TTTCTAGAGAGC-CCCTGACAAATCTGAGGCTTCATCAGAGAGCTAGAGCCATTTAACTT 858
QY 5294 TCCCTGTTTCAAGCTGCTTTGATTTTGCCCTTGATGCGCTCGGTATTTCTAACTT 5353
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 857 TCCCTTTTCAAGACCAACTCAATTTTCCCTTGATAC----- 818
QY 5354 TCCCTGTTTCAAGCTGCTTTGATTTTGCCCTTGATGCGCTCGGTATTTGCGATTTA 5413
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 817 -----CATTAATTTCTGATTTA 800
QY 5414 AGGTCTGCGCTTCACACTGGAAC-CAATTTGACCACTAGCTTCACTTCAACT 5472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 799 AGGTCTTAATTTTAACTGGAACCTGGAAGTGGCAATPACTAGTTCTGCTTGAAACT 740
QY 5473 GGAGATCTCTCTTATGATTTTATGATGTTTATGATGTTTATGATGTTTATGATGTTTAT 5532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 739 GGAGATCATTTCTTACATATTTATGATGTTTATGATGTTTATGATGTTTATGATGTTTAT 683
QY 5533 TGTATATTTTTCATAAGCTT-----AAGCAGATGCTTTTGTATTA 5575
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 TGTATATTTTTCATAAGCTT-----AAGCAGATGCTTTTGTATTA 623
QY 5576 TGCAATATATTAACGGGTGTGACGACCATAGCGACTTGTGAAGCTCCMAAGCTTCAACTG 5635
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 CACAGTATATGATGGGGATGTGCAATATGTCAGAGCTTGGGAGCTTTAAGCCTCAGTTA 563
QY 5636 TAACCTGAGCAAAACAGTAACTTC-----CTGGCAAGAGACAAAGTCTTTT 5686
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 TATAACCAACAAAAACAGAGCTCTAGATGTAACATTCCTGATCAAGTAACTTT 503
QY 5687 TAAAGTTACTGATGCTAGATCTGTGGGCTTCTAGTCCCTGGAAGGTGTTTCT 5746
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 TAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443
QY 5747 ATGCACAGCAGCTCAGAAATAAAAAAGCCATTTGAAACATCAGAGATGCCAAATTT 5806
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 ATTACAGCAGAGTGTCTAAACTTAAAAAGACTTTGAAACATCAGAGATGTTATTTTC 383
QY 5807 AACATGATTTTTCCTCCCTTTTGTCTAATCCAGTCCAGGTGGAAGAAAGTCTCTCTG 5866
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 ATTGGAATTT-----TTCTTCTAACCCAGTGGAGTTAGAAAGATTTATTTCT 330
QY 5867 TGTGAGA-----TTAAGCCCTGTCTTAATGATGATGACAAATGAT 5909
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 329 GGTAGCAATTTAATTTTATCAATCTTTTCTTACTTGTATGATGTTGACCGATTAAGT 270
QY 5910 GTGCTTAAGGCCATAGA-----TGTTTCTTAATGCAAGAAATCTGT 5953
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 269 GTGCTTAATCTGAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 210
QY 5954 TGTACGTTTTTTTGTGATGTTACTTTTATGCTGACCGAATTCATATCAGATGCAAGTG 6013
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 209 TGTAG--TTTTGATTTACTCTTATATGCTGACCTCATTTCAACATGCGATGAATA 152
QY 6014 AGTCTGTTCTTACAGATGATTTTGAATTTGATGATGATGATGATGATGATGATGATGATGAT 6073
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 151 AGTGAGTCTTTTACAAATGATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 92
QY 6074 TGCCCTCTTTTAAAGCAATGTCATATGTTCTTTTGTGATTAATTTGATTTGACAA 6133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91 TGCCATTTCTTTTAAAGCAATGTTGCAACATTTGATTTGATTAAGTTGTGATTTGACGA 32
QY 6134 CTGATTTTAAATTA 6147
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 31 CTGATTTAAATTA 18
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 15

US-60-324-185-30580
; Sequence 30580, Application US/60324185
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS IDENTIFIED THEREBY
; FILE REFERENCE: GX-0019-1 P
; CURRENT APPLICATION NUMBER: US/60/324,185
; NUMBER OF SEQ ID NOS: 35862
; SOFTWARE: PERL Program
; SEQ ID NO 30580
; LENGTH: 6632
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 445786.20
US-60-324-185-30580

Query Match 35.3%; Score 2172.8; DB 76; Length 6632;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 4220; Conservative 0; Mismatches 1797; Indels 539; Gaps 43;

QY	3	GGCGCGGGAGTAGAAGCCACTGAGCCATGGCGCAGGCAAGTTCACCCGAGCAGCGG	62
DB	190	GGCTAGCGCGGAGAGTAGTCGGAGGAGCCATGGCGCGGGAGCTCCACCGAGCAGCGC	249
QY	63	AGCCCGAGCAGCGG-----GGCGGAGGACACGCGGAGCAGAGTGTGTCTAGTGGCCAT	119
DB	250	AGCCCGAGCAGCGCGCGGAGGAGGAGTCCACGCGGCTGAGCCGAGCAGCGCGGCGGC	309
QY	120	GGGCCC-----GCAGCTGAAGCTCGGGAGCAGCTGGAGACCCCGCCGACGCGGA	169
DB	310	GGCCCTCGGCGAGCGCGGCGGACACACGCGCGGACCCCGCCATCGCTCGCTCGGA	369
QY	170	CCCGCGCACCAAGCTCCACAGAGAAATGGCCAGCTGTCTTCTCAACGGCGTAGCTGA	229
DB	370	CCCGCGCACCAAGCTCTACAGAGAAATGGTCACTGTCCACCATCAATCGCTAGCTGA	429
QY	230	ACAAGAGATGTCTATGTCACAGAGAAACCCAGGAGG-----	268
DB	430	GCAAGATGAGCTCAGCTCCAGGAGGAGTACCTAAATGGCCAGAAAGGAGCCCTGAACGG	489
QY	269	-----GCAGGAGGAGAGTGTGTGTATGATGAGGATGTTGACAGCG	307
DB	490	TCAGGAGCCCTAAACAGCGCAGGAGGAGAGAGTCAITTTGTCAAGAGGTTGGACAGAG	549
QY	308	AGAGTCAGAAAGATGTGAGAGAAAGACCCGAGTTGAAGAAATGGCGCCAACTCCACAGC	367
DB	550	AGACTCTGAAGATGTGAGCAAAAGAGACTCCGATAAAGAGATGGCTACTAAGTCAGCGGT	609
QY	368	TGTTGAAGATATCAAAAGATGGCGAGGAGGACATCAGAAATATTAAGTGAACAGATCCC	427
DB	610	TGTTCAAGATATCAGAGATGAGGAGGAGGAGACACCCGCAATAATTCGAACAGATTC	669
QY	428	TGCTTCAGAAACAAATGTGAAGAAATGGTACAGCTGTGAGTCCCGAGCTAATGATGT	487
DB	670	TTCTTCAGAAAGCAATTTAGAGAGCTTAAACAAACCCAGCTGAGTCCCGAGCTAATGATAT	729
QY	488	TGGCTTCAAGAAAGATATTTAAATTTTGTGTTTAAATTCACCGTGAAGAGGATAAAA	547
DB	730	TGGATTAAAGAGGTGTTAAGTTTGTGCTTTAAATTCACCTGTGAAAAGGATAGAC	789
QY	548	TGAAAGTCAAGTCTCTCAACTACTCTCACTGTCAAGAGGATGAAGCGGAGGCGGAG	607
DB	790	AGAGAGCCCTGACACTGTCAGCTACTCTCACTGTGAAGAAAGATGAAGGGGAGGAGCAGC	849
QY	608	AGCCTCTGTGCGAGCTGGAGACACAGGAGCCAGTGTGGAGACTGCGCTCGGAGATC	667

DB	850	A-----GGGGCTGGGACCAACAGGACCCAGCCTTGGGGCT-----GGAGAGC	894
QY	668	AGCATCCAAAGAAAGTAGCTGAAGCAATCCACAGAGAAAGCAAGAGCCCTGAAGCA	727
DB	895	AGCATCCAAAGAAAGCGAACCCCAACAAATCTACAGAGAAACCCGAAGAGACCTGAAGCG	954
QY	728	AGAACAGAGCAGCAGCAAAATCCCTTCAAGCCGAATCTGATCAAGCGGCTCAGGAGA	787
DB	955	TGAGCAAGGCGCACGCAAAATTTCTCCCGCAGCCGAATCTGGCCAA---GCATGGAGGA	1011
QY	788	AGCCAAAGATGAAGGAGAGAAACCAAGAGAAAGAGCCACCAAGTCCCCAGAAATCCCC	847
DB	1012	ATCCAAAGAGGAGAGAGAGAAACAGAAAGAACCTAGCAAGTCTTCAGAAATCTCC	1071
QY	848	GAGCAGCCAGTCAACAGAGTGAACAATCTCTTCAAGAAAGTCTTCACTCACGGTTG	907
DB	1072	GACTAGTCCCGTGACCGATGAACAGGATCAACCTTTCAAAAAATTTCTCACTCAAGGTTG	1131
QY	908	GGCGGCTGGCGCAAGAACAGCTTCAAGAAATCAAAAGAGGATGATCTGAAAATCTGC	967
DB	1132	GGCGGCTGGCGCAAAAGACCAAGTTTCAGGAAGCCGGAAGGAGGATGAAGTGAAGCTTC	1191
QY	968	CGAAGAGAGAGGAGCAGAGGACAGAAAGAGTAGACGAGAGAAAGAGAAAGAGACAGA	1027
DB	1192	AG	1251
QY	1028	GCAGGCTCGGAGGA-----GCAGGAGCGCGCAGAGGA	1060
DB	1252	GGTTGCTCCGAGAAACTGACCGCTCCGAGCAAGCCACACAGGAGCGCGCAGAGAG	1311
QY	1061	CACAGACAGCGCAGGTTGTACAGAGCTACAGAGAGGTTGAGAGTGTCTTGGAAAGACCA	1120
DB	1312	TGCCCCAGAGCCCCGCTTATCAGCTGAATATGAGAAAGTTGAGCTGCTCTCAGAGGAGCA	1371
QY	1121	GGTTGTGTACCTGGAGGACATCGTCAGAGAGAGTGTGTCTTGGACGAGCAAGTGT	1180
DB	1372	AGTCACTGGCTCGCAGGAGCTTCTGAAGAGAAACCTGTCTCGTTGGCGACAGAGTGT	1431
QY	1181	TGATGAGAGATGGAAGCCACCAAGAA---GTTGTTGAGAGGTCCACGTGAGACCGT	1237
DB	1432	TGATGAGAAATAGAGTCCACAGAGAGGTTGTGGCGAGTCCAGCTCAGACCGT	1491
QY	1238	GGAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1297
DB	1492	GGAGGAGAGAACCCGAGAGCAGAAA-----ACGAGAGTGA	1527
QY	1298	AGGAGAGAGAGTCTTGGCCCCCTCAGAAATCTGGCTGAGCCCGCAGGAGTCCCGCAGGA	1357
DB	1528	AGAAACAGCAGGGTCTGTGCCAGCTGAAGAAATTTGGTTGAAATGGATGCAAGAACCTCAGGA	1587
QY	1358	AGCTGAGCCTGTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAGGAGACCA	1417
DB	1588	AGCGAACTGCCAAGAGCTGTTGAAGCTCAAGAGAAACGTGTGTTTCCGAGAGGAGCC	1647
QY	1418	CATCAACTGACAGACCTAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1477
DB	1648	TACACAGGAGAGTACCTCAGTCTGATGAGAGAGTGTGTCTTCAAAACCCCGCAGGCGT	1707
QY	1478	TGTCAGTGAAGTGGAGATGTGTCTCTCAGGAGAGAAATCAGGATCAGGAGAGTCCCTT	1537
DB	1708	TGTGAGTGAAGTGAAGTGTGTCTATCAAGGAGAGAAATGAAGGTGAGGAGAGTCCACT	1767
QY	1538	GAAGAACTCTTTCAGTGAAGCTTAAAGAGCTGTCTGGGAAAGAGCAGAGAGGAGAA	1597
DB	1768	AAAGAGCTTTTACAGCACTGGCTTAAAGAGCTTTCTGGAAAGAGAAACAGAGAGGAA	1827
QY	1598	ACGAGAGTGGGGAG	1657
DB	1828	AAGAGGA---GGAGGAG	1884
QY	1658	AGAGAGTGTGATGAGCAGAGAGAGAGAGAGTCTTGGCTGTCTCCCGAGGAGGAGGAGG	1717
DB	1885	GGACAGCAGGAGGAGCAAAAGGGGAGAGAGTCTGTGCTCATCTCCCTGAGGAGCCGAGGA	1944

Qy	1718	GATAGCTGCTGTGAGAAAAGGGCCGCTGGACACACCCAGATGGGGAAGCTGTGAGAAAG	1777
Db	1945	GATCAAGTGTCTTGAAAAGGGCTTAACTGCGAGGTGTGACGATGTGGGAACTGTAAAGAAAG	2004
Qy	1778	AACTACTCTCCGATGTGAGAGAAAGAAAGAGAGAGATCACTCCCTGGGCACTCTTCAAAAA	1837
Db	2005	AGCTACTCTCGATGTGAGAGAAAAGAAAAGAAAGATGTCACTCCCTGGGCACTTCAAAAA	2064
Qy	1838	GATGTGTGACACCCAGAAAGGGGTCCGAAGACCTTCTGAGATGTGCAAGAGAAAGAGCT	1897
Db	2065	GATGTGTGACCCCAAGAAAGCGTGTGTGAAGGGCTTTCGAAATGTGTTAAAGAAATGTAGCT	2124
Qy	1898	GGAAAGGTCAAGAGGGCCACTTGTACTTCCATCTGTATGACACAGTGTCAAGAAATGCAGA	1957
Db	2125	GGACAAAGTCAAGAGCGCTACTTGTCTTCCACCGAGACACAGCTCTGAAAATGTACAGA	2184
Qy	1958	TGAAGTCAAAACTGTGTGTGAGAACAAAGCCAGAGAACCAAGCCGTAGGGTGTATAC	2017
Db	2185	AGAAATGAAGGAGCGCTGTGAAGGCCAAAGCCGAGAAACCAAGCCCAAGTGTGATAC	2244
Qy	2018	TTTCAGTGTCTTTGGAGACCTGATTTGTGTCCGATCATCAAGAAAGACAGAAAGCC	2077
Db	2245	CTCAGTATCTTTGGAGACTTAAATTTGTGTGGATCATCAAGAAAGACAGAAAGAG	2304
Qy	2078	ATCTCTTTCAGATGTATGAAGAGGGCCAAAGACACTCGGAGGGGACAGTCAAGAGCAGA	2137
Db	2305	GTCCTCTTCTGATGTAGGAAGGGGACCCAAAGCAATGTGGAGAGACACACGAAGAGCTGA	2364
Qy	2138	GGAGGCCAGCAAAAGCAAAAGAGCCGGAACAGACGCTGTCTCTGACAGACCCAGAGACA	2197
Db	2365	TGAGGCCGGAAGAAAGCAAAAGAGACGGGGACAGCGGGAATCTTGTCTGTGTCCAGAAACA	2424
Qy	2198	GGACCAAGCCGAAAGAAATTCTTACCCAGCCAGCGGAGAACCTCTTCCAAAGGGGAAG	2257
Db	2425	TGATCAAGCGCAGGAGAGTTCTCTCCCGAGCAAGCTGGAACCTCTACGAAGGGGAGGG	2484
Qy	2258	TGTCTTCCACTTTGGAGTCACTTTAAAGATTATGTCACTCCAAAGAAAATAATCCAAATCCAA	2317
Db	2485	CGTTTCCACTCTGGGAGTCACTTTAAAGATTATGTCAAGCCCAAGAAAATAATCCAAATCCAA	2544
Qy	2318	ACTGGAAGAAAGCCCGAAGAC-----TCTAGTGTAGAGCAAGTTGTCCACTGA	2365
Db	2545	GCTTGAAAGAAAGCCCAAGACTTCACTGCTGGGTCTGGTGTATGAACATTCCACTCCAGA	2604
Qy	2366	GATCCAAACCGATAGAGAGAAATCTTGGTTCATTAAAGAAATTCATCCCGGACGGCG	2425
Db	2605	CACGTAAACCGGTAAAGAAAGATCTGTGGTCTCAATCAAGAAATTATCTCGGACGAAG	2664
Qy	2426	GAAAGAAAAGGCGACAGCGGAAGCAAGAACCACTGTGTGAAGACTCAGGGCCAGTGTGA	2485
Db	2665	GAAAGAAAAGGCGCAATGTGTGAACCAAGAAACAAAGCCCTGTTGAAGAGCAGGGCCAAACAG	2744
Qy	2486	GATTAATGAGAGACACCTTAATGTGCCAGCGCTGTGCCCTGTCTGTGATTAATATGCAGT	2545
Db	2725	GGCCACCAAGATATGATCTGTATGTCTCCGCGCGTGTGTCCTCTGTCTGAATATGATGTCT	2784
Qy	2546	GGAGAGGGAGAAAGATGTGAAGCCCGAG-----GGGAATACGAGCTGCCAGCTGC	2595
Db	2785	AGAAAAGGAGAAATGTGAGGCGACAGCAAGCCCAAAAAGCCGACAGGACGCCGAGCAGA	2844
Qy	2596	TGGGGGCTGTGTACGTGTCCGAGAGCTCAGTATGAAGCTCTGTGTCCACTGTGTGATGTG	2655
Db	2845	AGGCAAGCCACTGAGGTGTCTCAAGAGAGCTCAGCAGAGTCAAGTTTCAATGATGTGACGAG	2904
Qy	2656	CAGTCAATTGATGTGACCCAGGGCACTCAACAGTGTCCAAAGGCGGTCTTCCCTGTGGATAT	2715
Db	2905	CTGTGTGTGACGGGACGAGGGCACTCACTTATTTGAAGAAAGTCTCTTCTTGGATAT	2966
Qy	2716	CGGTTCCGTAAACAGAACTCTTGAACACACAGCGGGAGAAAGCATGCCACTGTTTGAAG	2775
Db	2965	CTGCTTCAATGTACAACTCTTTGAACAGATGAAGCTGAAGCCGACATGTTAATCTGAGG	3022

Oy	2776	AGGCACTGAAAAAGACATCATTTGG---AGAAAGAACTCCCTGTGCTACCCACACGTTAC	2832
Db	3025	AGGTAATTGGAAAGAAAGTAATTGACAGAAAGAAACCCCCACCGTTACTGAACCTCTCG	3084
Oy	2833	CAGAGGGTAAAGATGCCATGACGACATGTCTACCACTGAAGTGGATTTCACCTCAGAA	2892
Db	3085	CAGAGAACAGAGAGGCCCGGGGCGACACGGTCTGTAGAGCGGAAATTCACCCCGAAG	3144
Oy	2893	CTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAAATTACCGAAGCATCCG	2952
Db	3145	CTGTGACAGCTGCAGAAACTCGAGGGCCATTGGGTGCGGAAGAAAGAACCGAAGCATCTG	3204
Oy	2953	GGGCGGAAGAGACACACAGCATGTGTCCGAGTTTCCAGCTACTGACCTCCACAGACA	3012
Db	3205	CTGCTGAAGAGACCAACAGAAATGTGTGTCAAGCTCTCCAGTTAAACGACTCCCGACACA	3264
Oy	3013	CCACAGAGAAAGCCACCCAGTTCCAGAGGTAGAGATGTGTGTCTAGATACAGAAAG	3072
Db	3265	CCACAGAGAGAGGCCACTCCGGTGCAGAGAGTGAAGGTGGCGTACTGACATGAAAGAC	3324
Oy	3073	AGGAGCCCAAGACCGCAGGCGCATCTCTCAAGCCGTTGCAGACAAAGTAAAGAGAGTCC	3132
Db	3325	AAGAGAGCCGGACTCAAGAGGCTCTCCAGGCAGTGGCAAAAAAGTAAAGAGAAATCC	3384
Oy	3133	AGGAGCCCTGACAC---CAGACTGTGCAGAGAACGGGGTCAAAAAGCACTGAGAGAGTTG	3189
Db	3385	AGCTGCTCGGCACCGGTGGGCGAAGATGTGTCTTCACTGTCTCAGAGACAGAGGACG	3444
Oy	3190	AGGAGGTAGAGAGAGACTCCGAAGTCTGTGGCTTCCGAGAAAGAGAGAGCGTTATGCCA	3249
Db	3445	AAAGACCAAGAAAGACAGCGCTGAAGCGTCCGGTCTAAGAAAGAGACGGAATGATTTGA	3504
Oy	3250	AAGGACCCGTGCAGAGAACTGAGCTGAGCATCTTGCAACAGGCGCTTGAGACTGGAACG	3309
Db	3505	AAGTAGATGCTCAGAGAGGCAAAAACCTGACCTTTTACAAAGGAGAGGTGTGTGGGAGA	3564
Oy	3310	CTACTCCAGAGAGCGTTG---AAGTCCAGAACTACCGGACAGATGTAGAC-----C	3357
Db	3565	CCACCCAGAAAGTTTGAAAAAGCTCTCAAGTACACAGAGCATAGATCTCCAGTACG	3624
Oy	3358	ATGTGCGCACGTGCCA-----GGTTATCAAGCTCCAGACGCTGATGG	3399
Db	3625	TTGTATACCATTTGTCAAGCCGAAACCTTAGCTGGGGTAAATCAACAGAGATGGATGG	3684
Oy	3400	AACAGGCGGTGGCCCTGAGTCAATCCGAAACCTTGACACAGACGTGAGACAAATGGAAGCA	3459
Db	3685	AACAGGCTATCCCCCTGACTCCGTGGAACCCCTTACACAGACGTGAGACTGATGGAAGCA	3744
Oy	3460	CTCCCTTAGAGAAATCAGACATGCGAGATGGGACACAGACAGATGAACCAATTGACGCC	3519
Db	3745	CCCCGTAGCCGAACTTGTGACGACACCGAGCAAAACCGAAGAACAGAGATTGTGAAATCC	3804
Oy	3520	AGGACAGTAAAGCCACTGACGTGTACGGCAGTCCAGGTCAACAGAAAGAGCGCGCTA	3579
Db	3805	ATGAGAGAAAGAGAGTGCATCTGTGATACCAAGTCAAGGGGGCACAAACAGAGGCAAGTTC	3864
Oy	3580	CTGCTCAGAAAGAGAGCCTTTCACACTTACTATATATTTCCAGCCACAGAGAAACATG	3639
Db	3865	CTGCACAGAAAGAGAGGCTCCACAGACCTTCCAGTTGTGTTCACAGAAAGAACTAAG	3924
Oy	3640	GGGAGAAACAGAGAAAGATGTTCTTGAACCTACACAGCAABAGCTTCTGTGACGCG	3699
Db	3925	AACATCTAAAGATGAAGACACTCTAGAGATACAGATTAAGAGGTGTCTCAGTGAACCTG	3984
Oy	3700	TGCCGCTTCTGGCAAAAGACTGAGGTGGGTCTCAAGAGGGTGA-----GGTTGACTGT	3750
Db	3985	TATTCATTCTCTCAAAAGACTGAGGGGACTGAAGGCTGACCAAGTATGCTGATGAGAAA	4044
Oy	3751	TGGATGAGAAAAAGTCMAAGAAACAGAGAGGTGTTGTATCACTCTGG-----ACCA	3804
Db	4045	CCAAAGAGTACCAATTTTGGAGAGACTTGAGGGGTCTATTAACACACAGGCAATTAACGTCA	4104
Oy	3805	ACATGCAAAAGGCTGCTGATGTGACATATACAGTAAAGTATGGATGGAGTGGCCGGGTCT	3866

Db 4105 GTCCGGAAGAGTCACTGAAGTTCCTTAAAGTGAAGGACAGAAGAGCTGAATGTA 4164
QY 3865 AGGAAAGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGAGG 3909
Db 4165 AAAAGGATGATGCTCTTGAACCTGCAGAGTCAAGTCTCCTCCATCCCTCCGCTGAGA 4224
QY 3910 GAGAGATGGAACCTGACGTTGAAAGAGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGT 3969
Db 4225 GAGAGATGGTAGTTCAAGTTCGAAAGGAGAGAAACAGAGAGCAGAGCCAACTCATGTGAATG 4284
QY 3970 AAGAAGG---TGAGCAGGAACACGCCGCTCTGAGCATGAAGGAACCTACGGGAAGCCAG 4026
Db 4285 AAGAGAGCTTGACACGAACAGCTGTACCGTATCTGAAGAGTCAAGTGAAGCAGCTCC 4344
QY 4027 TCCGACACTTGAATGCCCGAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCTTGGAGGAA 4086
Db 4345 TCCAGACAGTGAATGTGCCATCATAGATGGGCAAGGAAGTCAAGCAGTTTGGAGGAA 4404
QY 4087 GCCCTTCTC---TCCAGACCCAGACAAAGCAGGTTGCATAGAGTTCAAGTTCAAAGCC 4143
Db 4405 GCCCTCTCCCTGCTAGGTCAGAGAGGAGGAGTATGCACCAAAATTCAGATTCCAGAGCT 4464
QY 4144 TGGACACAAAGTCACTCAAAACAGCAGAGCTGTGGAAGA-----GGTCATGAAGACGG 4197
Db 4465 CTGAGGATCATTCCTTAACAGCGGCTGCAGAGGAGGAAGGCTCTTAGGAGAAACTG 4524
QY 4198 TTGTGATTTCCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGCACACTATTATCCAGCTG 4257
Db 4525 CCAACATTTTAGAAACAGGTGAAACGTTGGAGCCTGCAGGTCGACATTTAGTTCTGGAAG 4584
QY 4258 AAGAGTCTCTGCAACCGGTGGCCACTGGACTCTTCAGCATGCAGAGACACGGTACCCC 4317
Db 4585 AGAAATCCTCTGAAGAAATGAAGACTTCCCGCTCATCCAGGGGAAAGATGCTGTGCCCA 4644
QY 4318 TGGGGCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACCTCCTGCTCCTGAAAGCA 4377
Db 4645 CAGGGCCCGACTGTCAGGCAAAATCGACACAGGTGATAGTATCTGCTACTACCAGAAAG 4704
QY 4378 CCTTACATCTCGACCTCAAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAG 4437
Db 4705 GCTTAAGTCCGACCTGGAAGGAGAGAAACCAATCACTCACTCACTGAAGTCAAGTCAAG 4764
QY 4438 AGACAAGCCAGATGCTGTGCTCTGATGCTGACGGCAAGGAGATGACGAATCGAAAG 4497
Db 4765 TCGATGACGAGGTGTCTTGCC-----AGGAGGTCAAAGTCAAGTCAAGTCAAGTCAAG 4818
QY 4498 TCCTCAAGGCTGAACCTGAGATCTGGAACCTTGAGAGTGAAGCAACAAGATTGTGCTGA 4557
Db 4819 ATTTAGAGCTGAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAAACTTGTCCAAA 4878
QY 4558 ACGTCAATTCAGACAGCCGTTGACAGATTGCAAGTAC-----AGAAACAGCCCGGAAACTC 4614
Db 4879 ACATCATCCAGACAGCCGTTGACAGATTGACGATGAGTGAAGAAACAGCCACCAATGT 4938
QY 4615 ATGCTTATGATTCAGACACCCAGTTTC-----CTGCATGAGGCTTTGAC 4658
Db 4939 TGACGTCTGAGTTTACAGACACAAGCTCACGTGATAAAAGCTGACAGCCAGGAGCTGGAC 4998
QY 4659 AGCAGGAGCCCAACAGATGCTGCACAAAATGAAAGATGCCAGATGAACACCCAGTG 4718
Db 4999 AGGAAACGGAAGAAAGGAGAGGAACTCAGAGGCTCTGACAGGATGAACACCAATTA 5058
QY 4719 CCGCAGCCCA-----GAGAGGACTTGGCAAGTCTGACCGTTCTG 4757
Db 5059 CTTACGCCAAAGAGGAGTCAAGTCAACCGCAGTGGGACCAAGCATTTCTGATATTTCCA 5118
QY 4758 GAGGCATGGGCTCAGCCTCGGAAATGCTTCCGCGCTTGCAGTTGAAAGCCCGGTTGCA 4817
Db 5119 AAGACATGAGTGAAGCCTCAGAAAAGACCATGACTGTTGAGGTAGAAGGTTCCACTGTA 5178
QY 4818 A----- 4818

Db 5179 ATGATCAGCAGCTGGAAGAGGTGCTCCTCCATCTGAGGAGAGGAGGTGGAGCTGAA 5238
QY 4819 -----AGTAAGCATTTGAGAAGCTGC 4838
Db 5239 CAAAGTCTGTGCAGAGAAGATGATGTTGTCATGCTTGTGTAGCAGAAAAGATAGAGAAGTCA 5298
QY 4839 CTCTCAACCCCAAGA--TCCAAAAGGAGCATGCTGCTGATGCGCCCTCAGCTCCAAGCT 4896
Db 5299 TAGTTGAACCCGAAAGAGATGAAAAAGGTGATGATGTTGATGACCTTGAAAAACCAAGACT 5358
QY 4897 TAGCCCAAGCAGAGGCGCAGTGCCTCTGGAACCTTAACCAAGAAATCCCCAGACACCAACG 4956
Db 5359 CAGCCCTGGCTGATCTATGATGCTCAGAGGCTTAACCAAGAGTCCCAAGATACAAATG 5418
QY 4957 GACAAAGCTAAACGAGGAGGCGATCCCCCAAAAGTTGAGTCCAGGAAGAAAGAAATGT 5016
Db 5419 GACCAAAACAAAAGAGAGAGGAGTGCACCAGGAAGTAGAATTCAGGAAGGAAGAAAGTGC 5478
QY 5017 CTACCAAGTCAGTCAAGAGAAACAGGCCCGCAGGACAGAGAGGACCTGCAGAGGCCAAAG 5076
Db 5479 ACAGTGAATCAGATAAAGCGATCACACCCCAAGCAGCAGGAGAGTTACAGAAAACAGAGA 5538
QY 5077 GAGACCTGGCAGAACTCTAAGATGTTAGTTG-----CTCATTTG 5115
Db 5539 GAGATCTGCAAGTCAAGACTTACAGATCTTAAACATCATGCGAGTTAACTCATTTG 5598
QY 5116 ACATCTGTAAGACAGAAATGTGAAAAACAAGTCAAGAACAA-----GATGCTGCTGTTGGG 5171
Db 5599 CTGTTTGGAAAGACAGAAATGTGAAGACAAGTAGTAGAAGAAAATGAATGCTGCTGCTGAG 5658
QY 5172 ACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGAGGCGCTGCCAATGATTT 5231
Db 5659 A-CTGAAGACCAAGTATTTTCAGA--ACTTTGAGAAATGAGAGCAGGACACATCAACTGATCT 5716
QY 5232 CCACCAGTAGAGCACCCCGACAATTTCTGAGGCTTCATCGGAGCTAGAGCCAGCTCAACA 5291
Db 5717 CATTTCTAGAGAGC--CCCTGACAACTCTGAGGCTTTCATCAGGAGCTAGAGCCATTTAA 5775
QY 5292 TTTCTCGTTTCAAGACTGCTTTGATTTGCCCTTTGATGCCCTGCTGCTGATTTCTTAACA 5351
Db 5776 TTTCTCTTTTCCAAGACCAACCTACAATTTTCCCTTTGATAAC----- 5817
QY 5352 TTTCTCGTTTCAAGACTGCTTTGATTTGCCCTTGTATGCCCTGCTGCTGATTTTCGGATT 5411
Db 5818 -----CATATAAATTTCTGATT 5833
QY 5412 TAAGGTCTGCGTTCTCAACCTCGAAC--CAATTTCTGCCATACCTAGTTCCACTTCTCAAA 5470
Db 5834 TAAGGTCTTAAATTTCTTAACTCGAACTGGAGTTGGCAATACCTAGTTCTGCTTTGAAA 5893
QY 5471 CTGAGCATCCTCTTTATGATTTATATGATGATGTTTATGATAGTCTCCTCCTGTAACCT 5530
Db 5894 CTGAGTATCATTTCTTTACATATTTATGATGTTTAAAGTAG---TCTCTCTGATCT 5950
QY 5531 ATTGATATTTTTTTCTTAACGTTTT-----AAGCACATGCTTTTGTAT 5573
Db 5951 ATTGATATTTTTTTCTTAATGTTTAAAGGAAATGTGAGGATACTACATGCTTTTGTAT 6010
QY 5574 TATGCAATATATAACGGGTGTGAGCCATAGCAGCTTTGAAAAGCTCCAAGCCTCAAC 5633
Db 6011 CACACAGTATATATGATGGGCAATGTCATAGTGCAGGCTTTGAGGAGCTTTAAAGCCTCAGT 6070
QY 5634 TGTAACTCGAGCAAAACAGATAACATTC-----CTGGCAAGAGAGACAAGTCTTT 5684
Db 6071 TATATAACCCACAAAAACAGAGCTCTTAGATGTAACATTCCTGATCAAGGTACAATTC 6130
QY 5685 TTTAAAGTTTATGATGCTTTAGATCTGTGGGCTTCTAGTCTCTGAAAGTGGTGTGTTTC 5744
Db 6131 TTTAAAAATTCACCTTAATGATGAGTCCATATTTAGTGGTACTCTGAAATGGTCACTTTC 6190
QY 5745 CTATGCACGAGCTCAGAAATAAAAAACCCATTTTGAACATCCAGGATGTCACCAATA 5804
Db 6191 CTATTACGGAGTGTGCTAATAAATCTAAAAAGCATTTTGAACACATACAGAAATGTTCTATTG 6250

```
OY 5805 TTACCATGATTTTTCCTCCCTTTTGGCAATCCAGTCCAGTGGAAAGTCTCCTC 5864
Db 6251 TCATTGGGAAATTT-----TTCTTCTAACCCAGTGGAGGTAGAAAGATTAATTT 6303
OY 5865 TGTGTCAGA-----TTAAGCCCTGCTCTTAATGATATGACAAATGA 5907
Db 6304 CTGCTAGCAAAATTAACTTACATCTTTTCTACTTGTATGCTTGTGTTGACCGATTA 6363
OY 5908 GTGTGCTTAAGGCCATGAGA-----TGTTCCTAATGCAGAAAGAAATCT 5951
Db 6364 GTGTGCTTAATCTGAGGCAAGTAGTAATATGTTTATATGTTATGAAAGAAATTT 6423
OY 5952 GTTGTACGTTTTTTGATTTGATCTCTTCTAATGCTGACCGAATTCATATGCAGATCGAAG 6011
Db 6424 GTTGTAAAG--TTTTGATTTTACTCTTAATATGCTGACTGCAATGCACATGCGATGAAA 6481
OY 6012 TGAGTCTGTTCTTTAAGATGATATTTTGTATGATATGAGATTGCTGTGTATATC 6071
Db 6482 TAAGTCAGGTTCTTTAACAATGATATTTGATAGATCTGATGTGTGTGTGCCATATT 6541
OY 6072 TGTGCCCCCTTCTTTAAGAACAAATGTCATTAATGTTCTTTGATTAATTTGATTTGAC 6131
Db 6542 TGTGCAATTCCTTTAAGAACAAATGTTGCAACACATTCATTTGATTAAGTTGTGATTTGAC 6601
OY 6132 AACTGATTTAAATAAA 6147
Db 6602 GACTGATTTAAATAAA 6617
```

Search completed: December 13, 2002, 00:00:20
Job time : 11603 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:19:40 ; Search time 391 Seconds
(without alignments)
11377.814 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cgggcccggggagtagaag.....aaataaacatttgactac 6160

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 416730 seqs, 361097816 residues

Total number of hits satisfying chosen parameters: 833460

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5204.2	84.5	5236	6	US-10-152-319A-2063
2	2144.4	34.8	6608	1	PCT-US02-04915-43
3	1994.8	32.4	5931	5	US-09-724-676-12735
4	1994.8	32.4	5931	5	US-09-724-676A-12735
5	1920.8	31.2	5773	5	US-09-724-676-12736
6	1920.8	31.2	5773	5	US-09-724-676A-12736
C 7	550	8.9	563	6	US-10-152-319A-1312
8	162.2	2.6	178	6	US-10-152-319A-2137
9	124.6	2.0	180	6	US-10-152-319A-1193
10	119	1.9	753	5	US-09-724-676-12734
11	119	1.9	753	5	US-09-724-676A-12734
12	71	1.2	575	6	US-10-203-138A-9167
13	71	1.2	1969	6	US-10-203-138A-4036
14	70.8	1.1	1926	6	US-10-194-046-3
C 15	70.8	1.1	8705	6	US-10-273-678-16
C 16	70.8	1.1	8705	6	US-10-291-230-14
C 17	70.8	1.1	8705	6	US-10-291-249-14
C 18	70.8	1.1	16080	5	US-09-724-566A-48
19	68.4	1.1	107	5	US-09-620-607B-815
20	66.4	1.1	3489	6	US-10-194-046-1
C 21	65.8	1.1	659158	5	US-09-771-208A-20
22	61	1.0	700	6	US-10-203-138A-5915
23	56.4	0.9	2550	5	US-09-724-676-41735
24	56.4	0.9	2550	5	US-09-724-676A-41735
25	56.4	0.9	2709	5	US-09-724-676-41736
26	56.4	0.9	2709	5	US-09-724-676A-41736

27	56.4	0.9	2805	5	US-09-724-676-41728	Sequence 41728, A
28	56.4	0.9	2805	5	US-09-724-676A-41728	Sequence 41728, A
29	56.4	0.9	2964	5	US-09-724-676-41729	Sequence 41729, A
30	56.4	0.9	2964	5	US-09-724-676A-41729	Sequence 41729, A
31	56.4	0.9	3382	5	US-09-724-676-41723	Sequence 41723, A
32	56.4	0.9	3382	5	US-09-724-676A-41723	Sequence 41723, A
33	56.4	0.9	3637	5	US-09-724-676-41717	Sequence 41717, A
34	56.4	0.9	3637	5	US-09-724-676A-41717	Sequence 41717, A
35	56.4	0.9	3817	5	US-09-724-676-41745	Sequence 41745, A
36	56.4	0.9	3817	5	US-09-724-676A-41745	Sequence 41745, A
37	56.4	0.9	4072	5	US-09-724-676-41732	Sequence 41732, A
38	56.4	0.9	4072	5	US-09-724-676A-41732	Sequence 41732, A
39	56.2	0.9	474	6	US-10-203-138A-3499	Sequence 3499, Ap
40	56.2	0.9	906	6	US-10-273-334-33	Sequence 33, Appl
41	55.4	0.9	2142	6	US-10-152-319A-1533	Sequence 1533, Ap
42	55	0.9	905	6	US-10-273-334-11	Sequence 11, Appl
43	55	0.9	907	6	US-10-273-334-9	Sequence 9, Appl
44	55	0.9	907	6	US-10-273-334-13	Sequence 13, Appl
45	55	0.9	907	6	US-10-273-334-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-10-152-319A-2063
; Sequence 2063, Application US/10152319A
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgins, Brandon
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2063
; LENGTH: 5236
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_057103
US-10-152-319A-2063

Query Match 84.5%; Score 5204.2; DB 6; Length 5236;
Best Local Similarity 99.7%; Pred.No. 0; Mismatches 13; Indels 1; Gaps 1;
Matches 5223; Conservative 0

QY	1	CGGCGCGGAGCTAGAGAGCCACTGAGCCATGCGCGCAGGCACTTCCACCGAGCAGC	60
Db	1	CGGGCCGGCGGGAGTGAAGAGCCACTGAGCCATGCGCGCAGGCACTTCCACCGAGCAGC	60
QY	61	GGAGACCCCGAGCAGCCGGCGGGGAGGAGACACGCGCGAGCGAGCTGGTGTCTCAGTGGCCATG	120
Db	61	GGAGCCCGGAGCAGCCGGCGGGGAGGAGCACGCGCGAGCGAGCTGGTGTCTCAGTGGCCATG	120
QY	121	GGCCCGCAGCTGAAAGCTTCGGGAGCAGCTGGAGACCCCGCGCAGCCGAGCCCGCCACA	180
Db	121	GGCCCGCAGCTGAAAGCTTCGGGAGCAGCTGGAGACCCCGCGCAGCCCGCCAGCA	180
QY	181	AGCTTCCACAGAGAAATGCGCAGCTGTCTTCTGTCAACCGGCTTACCTGAACAGAGATG	240
Db	181	AGCTTCCACAGAGAAATGCGCAGCTGTCTTCTGTCAACCGGCTTACCTGAACAGAGATG	240
QY	241	TTCCATGTCCAAAGAGAAACACAGAGGGCGGAGGAGAAAGTCGTATGATGGATGTG	300
Db	241	TTCCATGTCCAAAGAGAAACACAGAGGGCGGAGGAGAAAGTCGTATGATGGATGTG	300
QY	301	GACAGCCGAGTCTCAGAGATGTGAGAGAAAAGCCGATTTAGAAAATGCGCGCACT	360
Db	301	GACAGCCGAGTCTCAGAGATGTGAGAGAAAAGCCGATTTAGAAAATGCGCGCACT	360
QY	361	CCACAGCTGTGAAGATATCAACAAAGATGGCGAGGAGAGACATCAGAAATTAATTGAC	420
Db	361	CCACAGCTGTGAAGATATCAACAAAGATGGCGAGGAGAGACATCAGAAATTAATTGAC	420
QY	421	AGATTCCTCTGCTTCAGAAAACAAATGTGGAGAAATGGTACAGCTCTGATCCACGCTA	480
Db	421	AGATTCCTCTGCTTCAGAAAACAAATGTGGAGAAATGGTACAGCTCTGATCCACGCTA	480
QY	481	ATGATGTGGCTTCAGAAAAGATTTAAATTTGGTTTAAATTCACGCTGAAGAG	540
Db	481	ATGATGTGGCTTCAGAAAAGATTTAAATTTGGTTTAAATTCACGCTGAAGAG	540
QY	541	ATTAATAATGAATAAGTCAGATATCTGTCCAACTACTCTGTCAAGAGATGAAGCGAG	600
Db	541	ATTAATAATGAATAAGTCAGATATCTGTCCAACTACTCTGTCAAGAGATGAAGCGAG	600
QY	601	GGGCGAAGCCTCTGTGGAGCTGGAGACCAACAGAGGCCAGTGTGAGACTGCGCTG	660
Db	601	GGGCGAAGCCTCTGTGTGGAGCTGGAGACCAACAGAGGCCAGTGTGAGACTGCGCTG	660
QY	661	GAGAGTCAGCATCCAAAAGAAAGTAGTGAAGCAATCCACAAGAAACAGAGGCAACC	720
Db	661	GAGAGTCAGCATCCAAAAGAAAGTAGTGAAGCAATCCACAAGAAACAGAGGCAACC	720
QY	721	TGAAGCAAGAACAGAGCAGCACAGAAATCCCTTCAGAGCCGATCTGATCAAGCGGCTG	780
Db	721	TGAAGCAAGAACAGAGCAGCACAGAAATCCCTTCAGAGCCGATCTGATCAAGCGGCTG	780
QY	781	AGGAAGAAAGCCAAAGATGAGAGAGAGAAAAACAAGAAAGAGCCACCAAGTCCCG	840
Db	781	AGGAAGAAAGCCAAAGATGAGAGAGAGAAAAACAAGAAAGAGCCACCAAGTCCCG	840
QY	841	AATCCCGAGAGAGCCAGTCAACAGTGAACAATCTTCTTCAAGAAATCTTCACTC	900
Db	841	AATCCCGAGAGAGCCAGTCAACAGTGAACAATCTTCTTCAAGAAATCTTCACTC	900
QY	901	ACGATTGGGCGCGCTGGCGCAGAGAGCCAGCTTCAAGAAATCAAAAAGAGATGATCTGG	960
Db	901	ACGATTGGGCGCGCTGGCGCAGAGAGCCAGCTTCAAGAAATCAAAAAGAGATGATCTGG	960
QY	961	AAACTGCGAGAGAGAAAGAGCACAAGGCAAAAAAGTGAATCCAGAGAAAGAAAGAAA	1020
Db	961	AAACTGCGAGAGAGAGAAAGAGCACAAGGCAAAAAAGTGAATCCAGAGAAAGAAAGAAA	1020
QY	1021	AGACAGAGCCAGCCTCGGAGAGAGAGAGCCGCGAGAAAGACACAGCCACGAGCTGTGT	1080
Db	1021	AGACAGAGCCAGCCTCGGAGAGAGAGAGCCGCGAGAAAGACACAGCCACGAGCTGTGT	1080
QY	1081	CAGCAGACTTACAGAGAGGTGAGCTGCTTTGGAAAGCCAGGTGGTGAATCTGAGGAGCAT	1140

Db	1081	CAGCAGACTACGAGAAAGGTGACCTGCTTTGGAAAGCCAGATTGTGACCTGGAGCAT	1140
Qy	1141	CGTCAGAGGAGAAAGTGTGCTCTTTGGCAACGGAAAGTGTTGATGAGAAATGGAAGCCC	1200
Db	1141	CGTCAGAGGAGAAAGTGTGCTCTTTGGCAACGGAAAGTGTTGATGAGAAATGGAAGCCC	1200
Qy	1201	ACCAAGAATTGTTGCAAGAGTCCACGTGAGCACCCTGAGAAAGACAGAGAGAGCAGG	1260
Db	1201	ACCAAGAATTGTTGCAAGAGTCCACGTGAGCACCCTGAGAAAGACAGAGAGAGCAGG	1260
Qy	1261	GAGAGAGAGGAGAGGCGTGAAGGGGGGCGTGTGTGAAAGGAAACAGAGAAATCTTGCCCC	1320
Db	1261	GAGAGAGAGGAGAGGCGTGAAGGGGGGCGTGTGTGAAAGGAAACAGAGAAATCTTGCCCC	1320
Qy	1321	CTGAGAAATCTGCGTGAAGCCCCCAGAGAGTCCCCCAGAGAGCTGAGCTGCTGAGAGCTGA	1380
Db	1321	CTGAGAAATCTGCGTGAAGCCCCCAGAGAGTCCCCCAGAGAGCTGAGCTGCTGAGAGCTGA	1380
Qy	1381	TGAAGACAGAGAGATGTGTCTCTGAGAGACCACTCACTCACTGACAGACTTAAGTC	1440
Db	1381	TGAAGACAGAGAGATGTGTCTCTGAGAGACCACTCACTCACTGACAGACTTAAGTC	1440
Qy	1441	CTGAAGAGAGAGACGCTGCCCCCAACCCCAAGAGGCAATGTCAGTGAGGTGAGATGCTGT	1500
Db	1441	CTGAAGAGAGAGACGCTGCCCCCAACCCCAAGAGGCAATGTCAGTGAGGTGAGATGCTGT	1500
Qy	1501	CCTCTCAGGAAAGATCAAGGTACAGGGAAGTCCCTTGAGAAATCTCTTCAGTACTCAG	1560
Db	1501	CCTCTCAGGAAAGATCAAGGTACAGGGAAGTCCCTTGAGAAATCTCTTCAGTACTCAG	1560
Qy	1561	GCTTAAAGAACCTGTCTGGGAAGACAGAGGGGAAACAGAGAGTGTGGGAGAGCAAG	1620
Db	1561	GCTTAAAGAACCTGTCTGGGAAGACAGAGGGGAAACAGAGAGTGTGGGAGAGCAAG	1620
Qy	1621	AGCCTGAGGAATACCAACAATTCAACCGAATCCCAAGAGTGTCTGATGAGAGAGG	1680
Db	1621	AGCCTGAGGAATACCAACAATTCAACCGAATCCCAAGAGTGTCTGATGAGAGAGG	1680
Qy	1681	GAGAGAGCTCTGCGTCCGCCAGAGACCTGAGGAGACCACTGTCTGGAGAAAGGCG	1740
Db	1681	GAGAGAGCTCTGCGTCCGCCAGAGACCTGAGGAGACCACTGTCTGGAGAAAGGCG	1740
Qy	1741	CGCTGGAAGCACTCCAGAGATGGGGAAGCTGAGGAAGAACTTACGATGAGAGAA	1800
Db	1741	CGCTGGAAGCACTCCAGAGATGGGGAAGCTGAGGAAGAACTTACGATGAGAGAA	1800
Qy	1801	AGAGAGAAAGGATACATCTCCTGGGACATCCTTCAAAAAGTGTGACACCAGAAACGGG	1860
Db	1801	AGAGAGAAAGGATACATCTCCTGGGACATCCTTCAAAAAGTGTGACACCAGAAACGGG	1860
Qy	1861	TCCGAAGACTTTCGAGAGTGAACAAGAGAGAGAGCTGAGAAAGGTCAAGAGCGCACCT	1920
Db	1861	TCCGAAGACTTTCGAGAGTGAACAAGAGAGAGAGCTGAGAAAGGTCAAGAGCGCACCT	1920
Qy	1921	TGTCCTCCACTGATAGCAGAGTCAAGAAATGCAAGATGAAGTCAAACTGTGGTAGG	1980
Db	1921	TGTCCTCCACTGATAGCAGAGTCAAGAAATGCAAGATGAAGTCAAACTGTGGTAGG	1980
Qy	1981	AACAAAGCCGAGAGAAACAAAGCGTGAAGGTGATCTTCAATGTCTTGGGAAGCATGA	2040
Db	1981	AACAAAGCCGAGAGAAACAAAGCGTGAAGGTGATCTTCAATGTCTTGGGAAGCATGA	2040
Qy	2041	TTTGTGTGATTCATCCAGAGAGAGCAAGAGGACATCTCTTCAGATGATGAAGAG	2100
Db	2041	TTTGTGTGATTCATCCAGAGAGAGCAAGAGGACATCTCTTCAGATGATGAAGAG	2100
Qy	2101	GGCCTAAGACACTGGAGGGGACAGTCAACAGACAGAGAGGCCACAGCAAGCAAGAG	2160
Db	2101	GGCCTAAGACACTGGAGGGGACAGTCAACAGACAGAGAGGCCACAGCAAGCAAGAG	2160
Qy	2161	CGGAAACAGACGCTGTTCTGTCCAGCACCCAGAGAGAGAACCAAGCGCAAGAAATTCTT	2220

Db 2161 CCGNACAGACGCTGTTCTCCAGCACCACAGGAGCAGGACCAAGCGCAAGAAAGTTCTCT 2220
Qy 2221 CACCCGACCCAGCGGGAAGCCCTTCCGAGGGGAAGGTGTCTCCACTTGGGAGTCATTTA 2280
Db 2221 CACCCGACCCAGCGGGAAGCCCTTCCGAGGGGAAGGTGTCTCCACTTGGGAGTCATTTA 2280
Qy 2281 AAAGATTAGTCACTCCCAAGAAAAAATCCAGTCAAACTGAAAGTGAAGAAAGCCGAAGACT 2340
Db 2281 AAAGATTAGTCACTCCCAAGAAAAAATCCAAAGTCAAACTGGAAGAGAAAGCCGAAGACT 2340
Qy 2341 CTAGTGTAGACAGTGTGTCCACTGAGATCGAAACCGAGTGAAGAAATCTTTGGGTTCCTCA 2400
Db 2341 CTAGTGTAGACAGTGTGTCCACTGAGATCGAAACCGAGTGAAGAAATCTTTGGGTTCCTCA 2400
Qy 2401 TTAAGAAATTCATCCCGGACCGGGAAGAAAGGGGCAGACGGGAAGCAAGAAACAGGCCA 2460
Db 2401 TTAAGAAATTCATCCCGGACCGGGAAGAAAGGGGCAGATGGGAAGCAAGAAACAGGCCA 2460
Qy 2461 CTGTGGAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCTTAATGTCCAGCCGCTCG 2520
Db 2461 CTGTGGAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCTTAATGTCCAGCCGCTCG 2520
Qy 2521 TGCCTCTGTCTGAGTATTAATGCAGTGGAGAGGAGAAAGATGGAAGCCAGGGGAATACGG 2580
Db 2521 TGCCTCTGTCTGAGTATGATGCGAGTGGAGAGGAGAAAGATGGAAGCCAGGGGAATACGG 2580
Qy 2581 AGTGTCCCCAGCTCTGTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTCC 2640
Db 2581 AGTGTCCCCAGCTCTGTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTCC 2640
Qy 2641 ACACGTGAGTGTGCGAGTCAATTGATGGGACCAAGGCGAGTCAACAGTGTGCGAAGCGGT 2700
Db 2641 ACACGTGAGTGTGCGAGTCAATTGATGGGACCAAGGCGAGTCAACAGTGTGCGAAGCGGT 2700
Qy 2701 CTCCTTGTGTGATATCCGCTTCCGTAAACAGAACTCTTGAACACACACGCGGGAGAGCCA 2760
Db 2701 CTCCTTGTGTGATATCCGCTTCCGTAAACAGAACTCTTGAACACACACGCGGGAGAGCCA 2760
Qy 2761 TGCCACCTGTTGAAGAGGTCACTGAAAGAACATCACTTGCAGAGAAACTCTCTGTGTCA 2820
Db 2761 TGCCACCTGTTGAAGAGGTCACTGAAAGAACATCACTTGCAGAGAAACTCTCTGTGTCA 2820
Qy 2821 CCCAGAGTTACAGAGGTTAAGATGCCATGACGATGGTCAACGTGTGCGAAGTGAAT 2880
Db 2821 CCCAGAGTTACAGAGGTTAAGATGCCATGACGATGGTCAACGTGTGCGAAGTGAAT 2880
Qy 2881 TCACCTCAGAGCTGTGACACGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAGTTA 2940
Db 2881 TCACCTCAGAGCTGTGACACGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAGTTA 2940
Qy 2941 CCGAAGCATCGGGGGCCGAAGAGACCACAGACATGGTGTCCGAGTGTTCACAGCTGACTG 3000
Db 2941 CCGAAGCATCGGGGGCCGAAGAGACCACAGACATGGTGTCCGAGTGTTCACAGCTGACTG 3000
Qy 3001 ACTCCCGACACACAGAGAGAACCCAGTTCAGGAGGTGAGAGTGGTGTGTAG 3060
Db 3001 ACTCCCGACACACAGAGAGAACCCAGTTCAGGAGGTGAGAGTGGTGTGTAG 3060
Qy 3061 ATACAGAGAGAGGAGCGCCAGACGCGAGCCATCTCCAGAGGCTTGCACACAGGTGA 3120
Db 3061 ATACAGAGAGAGGAGCGCCAGACGCGAGCCATCTCCAGAGGCTTGCACACAGGTGA 3120
Qy 3121 AAGAGAGTCCAGGTGTCTCAACCCAGACTGTGCGAGAGAACGGGTCAAAAGCACTGG 3180
Db 3121 AAGAGAGTCCAGGTGTCTCAACCCAGACTGTGCGAGAGAACGGGTCAAAAGCACTGG 3180
Qy 3181 AGAAGTTGAGAGGTAGAGAGGACTCCGAAGTGTCTGGTTCGGAGAAAGAGAGGACG 3240
Db 3181 AGAAGTTGAGAGGTAGAGAGGACTCCGAAGTGTCTGGTTCGGAGAAAGAGAGGACG 3240
Qy 3241 TTATGCCGAAAGGACCGGTGACAGAGCTGAGCTGAGCATCTTGTGCAAGGGCTCTGAGA 3300
Db 3241 TTATGCCGAAAGGACCGGTGACAGAGCTGAGCTGAGCATCTTGTGCAAGGGCTCTGAGA 3300

Qy 3301 CTGACAGGCTACTCTCAGAGAGCCTTGAAGTTCCTGAAGTCACGGCAGATGTAGACCATG 3360
Db 3301 CTGACAGGCTACTCTCAGAGAGCCTTGAAGTTCCTGAAGTCACGGCAGATGTAGACCATG 3360
Qy 3361 TCGCCAGTCCGAGGTATCAAGCTCCAGCAGCTGATGGAACAGGCCGTGGCCCTCAGT 3420
Db 3361 TCGCCAGTCCGAGGTATCAAGCTCCAGCAGCTGATGGAACAGGCCGTGGCCCTCAGT 3420
Qy 3421 CATCCGAAACCTTGCACAGACAGTGAGACAAATGGAAGCACTCCCTTAGCAGATTCAGACA 3480
Db 3421 CATCCGAAACCTTGCACAGACAGTGAGACAAATGGAAGCACTCCCTTAGCAGATTCAGACA 3480
Qy 3481 CTGAGATGGGACACACAAAGATGAAACCAATTGACGCCAGGACAGTAAGCCACTGCAG 3540
Db 3481 CTGAGATGGGACACACAAAGATGAAACCAATTGACGCCAGGACAGTAAGCCACTGCAG 3540
Qy 3541 CTGTCCAGCAGTCAACAGTCAACAGAGAGAGGGCGGTACTGTCTCAGAAAGAGGCGCTT 3600
Db 3541 CTGTCCAGCAGTCAACAGTCAACAGAGAGAGGGCGGTACTGTCTCAGAAAGAGGCGCTT 3600
Qy 3601 CGACACTTACCTTAATATGTTCCAGCCCCAGGAAGAAATGCGGGAAGAACACAGGAAGAGATG 3660
Db 3601 CGACACTTACCTTAATATGTTCCAGCCCCAGGAAGAAATGCGGGAAGAACACAGGAAGAGATG 3660
Qy 3661 TTCTTGAACCTTACACAGCAAGAGCTTACTGCTGCAGCCGTGCCCGTCTGGCAAAAGACTG 3720
Db 3661 TTCTTGAACCTTACACAGCAAGAGCTTACTGCTGCAGCCGTGCCCGTCTGGCAAAAGACTG 3720
Qy 3721 AGTGGGTCAAGAGGTGAGGTTGACTGGTGGATGGAGAAAAGTCAAAAGAGAACAGG 3780
Db 3721 AGTGGGTCAAGAGGTGAGGTTGACTGGTGGATGGAGAAAAGTCAAAAGAGAACAGG 3780
Qy 3781 AGGTGTTTGTACACTCTGGACCCAAACAGTCAAAAGGCTGTGTGATGACATATCACAGTG 3840
Db 3781 AGGTGTTTGTACACTCTGGACCCAAACAGTCAAAAGGCTGTGTGATGACATATCACAGTG 3840
Qy 3841 AAGTGTGGAGTGGCGGGTGTGAGGAAAAGGAGAGTACTGAAAGTGCAGAGTCTTAGCC 3900
Db 3841 AAGTGTGGAGTGGCGGGTGTGAGGAAAAGGAGAGTACTGAAAGTGCAGAGTCTTAGCC 3900
Qy 3901 TGGAGGAGGAGAGATGGAACCTGACGTTGAAAGGAGAAAAGGAGACAAAGCCAGC 3960
Db 3901 TGGAGGAGGAGAGATGGAACCTGACGTTGAAAGGAGAAAAGGAGACAAAGCCAGC 3960
Qy 3961 AAGTGTGAGGAGGAGTGAAGGAGAAACAGCCCTCTGAGACATGAAGGAACCTTACCGGA 4020
Db 3961 AAGTGTGAGGAGGAGTGAAGGAGAAACAGCCCTCTGAGACATGAAGGAACCTTACCGGA 4020
Qy 4021 AGCCAGTCTCTGACACTTTGACATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCTTG 4080
Db 4021 AGCCAGTCTCTGACACTTTGACATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCTTG 4080
Qy 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGAGTGTGATAGAGTTCAAGTTCAAA 4140
Db 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGAGTGTGATAGAGTTCAAGTTCAAA 4140
Qy 4141 GCCTGGACACAAACAGTCACTCAACACAGCAAGCTGTGGAAAAGGTATAGAAACGGTTG 4200
Db 4141 GCCTGGACACAAACAGTCACTCAACACAGCAAGCTGTGGAAAAGGTATAGAAACGGTTG 4200
Qy 4201 TGATTTTCAGAGACAGGTGAAGTCCAGAGTGTGTAGGTGCACACTTATTACCAGCTGAGA 4260
Db 4201 TGATTTTCAGAGACAGGTGAAGTCCAGAGTGTGTAGGTGCACACTTATTACCAGCTGAGA 4260
Qy 4261 AGTCTCTGTCAAACGGGTGGCCACTGGACTCTTTCAGCATGAGAGGACACGGTACCCCTGG 4320
Db 4261 AGTCTCTGTCAAACGGGTGGCCACTGGACTCTTTCAGCATGAGAGGACACGGTACCCCTGG 4320
Qy 4321 GGCTGTAGTCTCAGGAGAGAAATCCCAATCATAGTAACTCTCTGCTCTTGAAGACACC 4380
Db 4321 GGCTGTAGTCTCAGGAGAGAAATCCCAATCATAGTAACTCTCTGCTCTTGAAGACACC 4380

	Query Match	34.8%	Score 2144.4	DB 1	Length 6608	
	Best Local Similarity	64.3%	Pred. No. 0			
	Matches 4215	Conservative 0	Mismatches 1796	Indels 543	Gaps 45	
QY	3	GGCGCGGGGAGTAGAAGACCACTGAGCCATGCGCGGAGGACGTTCCACCGAGACGCG	62			
DB	162	GGCTAGAGGCGGGAGAAATGCGAGAGACCCTATGGCGCGCGGAGGCTCCACCGAGACGCG	221			
QY	63	AGCCCCGAGCAGCCG---GCGGGGAGCGACACGCGGAGCGAGCTGTGCTCAGTGGCCAT	119			
DB	222	AGCCCGAGGACGCGCCCGGAGGAGCTCCACGCGCTGAGCCCGAGCCAGCCGCGGCG	281			
QY	120	GGGCGCCGAGCTGGAAGCTCTGGGA-----GCAGCTGAGAACCCCGCGACGCGGAC	170			
DB	282	GGCCCTCGGCGGAGGCGGCGCGACAGACACACCGCGGACCCCGCATGCTCTCGGAC	341			
QY	171	CCCCCACCAGGCTCCACAGAAATGCGCCAGCTGTCTTCTGTCAAGCGGTAGCTGA	230			
DB	342	CCCGCACCAAGCTCTTACAGAAATGCTCAGCTGTCCACCATCAATGCGGTAGCTGAG	401			
QY	231	CAAGAGATGTCATGTGTCAAGAGGAAAAACGAGAGG-----	268			
DB	402	CAAGATGAGCTCAGCTCTCAGAGGGGTACCTTAATGCGCAGAAAGAGCCCTGAACGT	461			
QY	269	-----GCAGAGGAGAAAGTCTGTGATGAGATGTTGACAGCGA	308			
DB	462	CAAGAGCCCTTAACAGCGCAGAGGAGAAAGATGATTTGTCAAGAGGTTGACAGAGA	521			
QY	309	GAGTCAGAAAGTGTGAGAGAAAAGACCGAGTTGAAGAAATGCGCGCACTCCACAGCT	368			
DB	522	GACTCTGAAGATGTGAGAGGAAAGACCTCCGATTAAGAGATGGCTACTAAGCTCAGCGGTT	581			
QY	369	GTTGAAGATATCAAAAGATGGGAGGAGGAGCATCAGAAATAATGAACAGATCCCT	428			
DB	582	GTTCAAGCATACAGATGATGGGAGGAGGA---ACCGAATATGAAACAGATCTCT	638			
QY	429	GCTTCAGAAAACAATGTGGAAGAAATGTTACAGCTGTGAGTCCCAAGCTATATGTT	488			
DB	639	TCTTCAGAAAGAAATTTGAAGAGCTAACACACCACATGAGTCCCAAGCTATATGATTT	698			
QY	489	GCGTTCAAGAAAGTATTTAAATTGTTGGTATTAAATTCACGGTGAAGAGATAAAAAT	548			
DB	699	GGATTTAAGAAAGTGTATTAAGTTGTGTGGCTTTAATTCATCTGTAAAAAAGATTAAGACA	758			
QY	549	GAAAGTCAGATCTGTCCCACTACTCATCTGTCAAGAAAGATGAAGGCGAAGGGCAGAA	608			
DB	759	GAGAAAGCTGACATGTCAGCTACTCATCTGTGAAGAAAGATGAAGGGGAGGAGCAGACA	818			
QY	609	GCGCTGTGCGAGCTGAGAGCAACAGAGAGCCAGTGTGAGAGACTGCGGTGAGAGTGA	668			
DB	819	-----GGGCGCTGCGGACCAACAGAGACCCCAAGCTTTGGGCGCT-----GAGAGACA	863			

QY 669 GCATCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAAAGCAAGGACCCCTGAAGCAA 728
Db 864 GCATCCAAAGAAAGCGAACCCAAACAATCTACAGAGAAACCCGAAGAGACCCCTGAAGCGT 923
QY 729 GAACAGAGCAGCAGAGAAATCCCCCTTCAAGCCGAATCTGATCAAGCGGCTGAGGAGAA 788
Db 924 GAGCAAGCCACGAGCAAAATTTCTCCCGCAGCCGATCTTGCCNAGCAGT---GGAGGAA 980
QY 789 GCCAAGATGAAGGAGAGAGAAACAAAGAGAAAGAGCCACCAAGTCCCCAGCAATCCCG 848
Db 981 TGCAAGAGGAGAGAGAGAGAAACAAAGAAAGAAACCACTAGCAAGTCTGCAGAAATCTCCG 1040
QY 849 AGAGCCCACTCAACAGTGAAGCAACATCTTCTTCAAGAGTCTTCACTCAGCGTTGG 908
Db 1041 ACTAGTCCCGTGACAGTGAACAGGATCAACCTTCAAAAATTTCTTCACTCAAGGTTGG 1100
QY 909 GCCGGCTGGCGCAAGAGACCCAGCTTCAAGAAATCAAAAGAGGATGATCTGGAACTGCC 968
Db 1101 GCCGGCTGGCGCAAAAGACCAAGTTTCAGNAGCCGAAAGGAGATGAATGGAAAGCTTCA 1160
QY 969 GAGAAGAAAGGAGCAGAGGCGCAGAAAAAGTAGACAGAGGAGAAAGAAAGACACAGAG 1028
Db 1161 GAGAAGAAAGGAGAAAGAGCCAGAGAAAAAGTAGACACAGAGAAAGACGGAAGGACAGAG 1220
QY 1029 CCAGCCTCGAGG-----GCGAGAGCCGCGCAGAGAC 1061
Db 1221 GTTGCCTCCGAGAAATGACCGCTCCGAGCAAGCCACCCACAGGAGCCGCGAGAAAGT 1280
QY 1062 ACAGACCAGGCGAGTTGTGACAGACATACGAGAAAGTGAGGCTGCCTTTTGGAGACAG 1121
Db 1281 GCCACAGCCCGGTTATCAGCTGAATATAGAAAGTTGAGCTGCCCTCAGAGAGCAAA 1340
QY 1122 GTTGGTCACTGGAGGATCGTCAGAGGAGAAAGTGTCTCTCTTTGGCAACGGAAGTGT 1181
Db 1341 GTCAGTGGCTCGCAGGACCTTCTGAAGAAACCTGCTCGTTGGCGACAGAGTGT 1400
QY 1182 GATGAGAAGATGGAGCCACCAAGAA---GTGTTGACAGAGTCCAGTGAGACCCTG 1238
Db 1401 GATGAGAAAAATAGAAATCCACCAAGAGAGGTTGTGCGCGAAAGTCCACGTCAGCACCGTG 1460
QY 1239 GAGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGCTGAAGGGGCGTGGGTAGAA 1298
Db 1461 GAGGAGAAACCGAAGACAGAA-----ACGGAGTGGAA 1496
QY 1299 GGAACAGGAGAACTCTTGCCTCCCTGAGAACTGGCTGAGCCCGCAGGAGTCCCCAGGAA 1358
Db 1497 GAACACAGGGTCTGTGCCAGCTGAAGAAATGTTGGAATGATGACAGAACTCAGGAA 1556
QY 1359 GCTGAGCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGAGGAGACCCAC 1418
Db 1557 GCCGAACCTGCAAGGAGCTGGTGAAGCTCAAGAAACGTTGTGTTCCGGGAGGACCT 1616
QY 1419 ACTCACTGACAGACCTAAGTCTTGAAGAGAGAGCTGCCCAACACCCAGAGGCAAT 1478
Db 1617 ACAGGAGGCTGACCTCAGTCTGATGAGAGGTTGTGTCAAACCCCGCGAAGGCGTT 1676
QY 1479 GTCAGTGAAGTGAAGTGTCTCTCAGGAAAGAAATCAAGGTACAGGGAAGTCCCTTG 1538
Db 1677 GTGAGTGAAGTGAAGTGTCTATCACAGGAGAGATGAAGTGCAGGAGTCTCACTA 1736
QY 1539 AAGAACTCTTCACTAGTCTAGGCTTAAAGAGCTGTCTGGGAGAAAGCAGAGGGGAAA 1598
Db 1737 AAGAAAGCTTTTACCAAGCTGGCTTAAAAAAGCTTCTCGAAGAAACAGAAAGGAAA 1796
QY 1599 CGAGGAGGTCGGGAGCAGAGAGCTGGAGAAATACCAACATTCACCGAATCCCA 1658
Db 1797 AGAGGA---GGAGGAGACGAGGAAATCAGGGAGCACACTCAGGTTCCAGCGATCTCCG 1853
QY 1659 GAGAGTCTGATGACAGAAAGGAGAGAGCTCTCGCTGCTCCCGCAGGAGCCCTGAGAG 1718
Db 1854 GACAGCAGGAGGAGCAAAAGGCGAGAGCTCTGCTCATCCCTGAGGAGCCCGAGGAG 1913

QY 1719 ACCACGTGTCTGAGAAAGGGCCGCTGGAAAGCAACCCAGGATGGGGAAGCTGAGGAAGGA 1778
Db 1914 ATCACGTGTCTGGAAGAGGGCTTTAGCCGAGGTGCAGCAGATGGGGAAGCTGGAAGAGGA 1973
QY 1779 ACTACTTCCGATCGAGAGAAAGAGAGAGGATCACTCCCTGGGCATCTCTTCAAAAAG 1838
Db 1974 GCTACTTCCGATCGAGAGAAAGAAAGAGAGGTGTCACTCCCTGGGCATCTTCAAAAAG 2033
QY 1839 ATGGTGACACCCCAAGAAACGGGTCCGAAGACCTTCTGAGAGTCAAGAGGAGGAGCTG 1898
Db 2034 ATGGTGAGCGCCCAAGAGCGTGTGTAGACGGCTTCGGAAGTGTATAAGAAAGATGAGCTG 2093
QY 1899 GAGAAAGTCAAGAGCGCCACCTTGTCTTCCACTGATAGCAGTGTGAGAAATGCAAGAT 1958
Db 2094 GACAAAGTCAAGAGCGCTACCTTGTCTTCCACCGAGACACAGCCTCTGAAATGCAAGAA 2153
QY 1959 GAAGTCAAAAACCTGTGTGAGGAAACAAAGCCAGAGAGCAAAAGCGTAGGGTGATCT 2018
Db 2154 GAATGAAAGGAGCGTGGNAGAGCCAAAGCCGAGNAGCAAAAGCCAGAGTGGATACC 2213
QY 2019 TCAGTGTCTTGGGAAGCAGCTGATTTGTGTGCGATCATCCAAAGAGAGAGCAAGGAAGCA 2078
Db 2214 TCAGTATCTTGGGAAGCTTTAAATTTGTGTGGATCATCCAAGAAAAAGAGCAAGGAAGG 2273
QY 2079 TCCTCTTCAGATGATGAAGAGGGCCNAGGACACTGGAGGGGACAGTTCACAGAGCAGAG 2138
Db 2274 TCCTCTTCTGTGAGGAAGGGGACCAAAAGCAATGGAGGAGACCAACCAAGAAAGCTGAT 2333
QY 2139 GAGGCAGCAAAAGACAAAGAACCGGAAACAGACGCTTCTCCAGAGACCCAGGAGCAG 2198
Db 2334 GAGCCGGAAGAAACAAAGAGACGGGACAGCGGATCTTCTGTTCCCAAGAAACAT 2393
QY 2199 GACCAAGCGCAAGAAAGTCTCTCACCCGAGCAGCGGAGACCCCTTCCGAAGGGGAAGGT 2258
Db 2394 GATCCAGGCGAGGAAAGTTCTCTCCCGAGCAAGCTGGAAGCCCTACCGAAGGGAGGC 2453
QY 2259 GTCTCACTTGGAGTCAATTTAAAGATTTAGTCACTCCAAAGAAAAAATCAAGTCAAAA 2318
Db 2454 GTTTCACCTGGGAGTCATTTAAAGGTTTAGTCACGCCAAGAAAAAATCAAAAGTCCAAG 2513
QY 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGAGTTGTCACCTGAG 2366
Db 2514 CTGGAAGAGAAAGCGAAGACTCCATAGTGGGTCTGGTGTAGAACATTCACCTCCAGAC 2573
QY 2367 ATCGAACCCGATAGAGAAGAAATCTTGGGTTTCCATTAAGAAATTCATCCCGGACGCGCG 2426
Db 2574 ACTGAACCCGTTAAAGAGATCTCTGGTCTCAATCAAGAGTTTATCTCGACGAGG 2633
QY 2427 AAGAAAGGCGACAGCGGAGAGCAAGAACCAAGCCTGTGGAAGACTCAGGGCCAGTGGAG 2486
Db 2634 AAGAAAGGCGCAGATGGGAAACAAAGAACAGCCCTGTTGAAGACGCGAGGCCAACAGGG 2693
QY 2487 ATAAATGAGGACACCTTAATGTCCAGCGCTGTCGCTCTGTCTGAGTATATGCAATG 2546
Db 2694 GCCAACGAAGATGACTGTATGTCCCGGCGTGGTCCCTCTGTCTGAGTATGATGTGTA 2753
QY 2547 GAGAGGAGAGAAATGG-----AAGCCCGAGGGAATACGGAGCTGCCCGCAGCTGCTG 2597
Db 2754 GAAAGGAGAAATGAGGCGACAGCAAGCCCAAAAGGCGCAGAGCAGCCCGGACAGAG 2813
QY 2598 GGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGGTCCACTGTGAGTGTGCA 2657
Db 2814 GCAGCCACTGAGGTGTCCAAGGAGCTCAGCGAGAGTCAAGGTTTATATGATGAGCAGCAGCT 2873
QY 2658 GTCAATGTAGGGACCGGCGAGTCAACAGTGTGGAAGCGGTCTCTCTGATGATATCC 2717
Db 2874 GTCGCTGAAGGAGCGAGGCGAGCTACCAATTTAAGAAAGGTTCTCTCTTGGATATCT 2933
QY 2718 GCTTCGTTAAACAGAACTCTTTGAACACAGCGGAGAGCCATGCCACCTGTTTGAAGAG 2777
Db 2934 GCTTCAGTGAAGAACTCTTTGAACAGTGAAGCTGAAGCCCGCAGCTGTTAACTGAGGAG 2993
QY 2778 GTCACCTGAAAAAGACATCATTCG---AGAAGAAACTCTGTGTCTCAGCCAGAGCTTACCA 2834

Db 2994 GTATTGGAAAGAAATATTGTCAGAAAGAAACCCCGCTTAAGTCACTTGCCCA 3053
Qy 2835 GAGGCTAAAGATGCCATGACGACATGTGTCAACAGTGAAGTGAATTTCACTCAGAACCT 2894
Db 3054 GAGAAACAGAGAGGCCCGGGGGGACACGGTCCGTATGTAGAGGGGAATTTGACCCCGAAGCT 3113
Qy 2895 GTGACAGCCACAGAGACTCAGAGGCTCCGTAATGAAGAAATTTACCGAGACATCGGGG 2954
Db 3114 GTGACAGCTGCAAGAACTGACAGGGCCATGTGGGTCCGAAAGAAAGCAACCGAATCTCTCT 3173
Qy 2955 GCGGAGAGACACAGACATGCTGTCCGAGTTTCCGAGTTCAGCTGACCTCCCGACAGACC 3014
Db 3174 GCTGAAGAGACACAGAAATGCTGACAGTCTCCAGTTTAACGACTCCCGACAGACC 3233
Qy 3015 ACAGAGAGACACCCCACTCAGAGGTAGAGAGTGTGTCTAGATACAGAAAGAG 3074
Db 3234 ACAGAGAGGCCACTCCGGTGCAGAGGTGGAAGGTGGGTACTGTACATGAAAGAGCA 3293
Qy 3075 GAGCGCCAGACGACAGGCATCTCTCAAGCCGTTGCAGACAGAGTGAAGAGAGTCCAG 3134
Db 3294 GAGAGGCGGACTCAAGAGTCTCTCCAGGAGTGCAGAGAAAGTGAAGAGAAATCCAG 3353
Qy 3135 GTGCGTGCACAC---CAGACTGTGACAGAAACGGGTCAAAAGCACTGAGAAAGTTGAG 3191
Db 3354 CTGCTTGACACCGTGGGCGAGAGATGTGCTTCAGCTGTGCAGAGACAGAGGACAGAA 3413
Qy 3192 GAGGTAGAGAGAGACTCCGAGTGTGCTTCGAGAAAGAGAGAGAGCTTAATGCGAAA 3251
Db 3414 AGACAGAGAGAGAGAGCTGAGAGCTGCGGTCTGAAGAAAGAGACGAGATGTATGTTAAA 3473
Qy 3252 GAGCCGTCAGAGAGCTGAGAGTGAAGTCTTGCAACAGGGCTCTGAGACTGAACAGCT 3311
Db 3474 GTAGATGCTCAGAGAGCAAAAAGTGAAGCTTTTACAGAGGAGAGTGTGGGACAGACC 3533
Qy 3312 ACTCAGAGAGCTTG---AAGTCTGTGAAGTCAAGGACAGATGTGAC-----CAT 3359
Db 3534 ACCCGAGAAAGCTTTGAAAAAGCTCTCAAGTCAAGAGAGATGATGCTTCAGTGAAGCTT 3593
Qy 3360 GTGCGCACAGTGCCA-----GGTTATCAAGCTCCAGCAGCTGATGNA 3401
Db 3534 GTAACCACTTGTCAGACCGGAAACCTTAGTGGGTAAATATCAGAGAGATGATGATGNA 3653
Qy 3402 CAGGCGGTGGCCCTGAGTCAATCCGAAACCTTGACAGACAGTGAACAAATGGAAGACT 3461
Db 3654 CAGGCTATCCCTCGACTCGGTGGAACCCCTACAGACAGTGAATGTATGGAAGCACCC 3713
Qy 3462 CCCTTAGCAGATTTCAGACACTGACATGAGGACACAGCAAGATGAACCAATTGACGCCAG 3521
Db 3714 CCGGTAGCCGACTTTTGACGACACAGGACCAACCCAGAAAGACGAGATTGTGAAAAATCCAT 3773
Qy 3522 GACAGTAAAGCCACTGACAGTGTCAAGCAGTCAAGTCAAGAGAGAGCGGCTACT 3581
Db 3774 GAGAGGAATGAGGTGCAATCTGTATCCAGTCAAGGGGGGACAGAGACAGAGTCTCT 3833
Qy 3582 GCTCAGAAAGAGAGCCTTTCACACTAATTAATGTTTCAAGCCAGAGAGAAACATGGG 3641
Db 3834 GCACAGAAAGAGAGGCTTCACACACTTTCAGTTTGTTCAGAGAAAGAACTAAAGNA 3893
Qy 3642 GAGAGAACAGAGAGATGTTTGAACCTTACAGAGAGAGACTTATGCTGCACGCCG 3701
Db 3894 CAATCAAGATGAGAGACACTTGAAGCATACAGATAAAGAGAGTGTCACTGGAACCTGTA 3953
Qy 3702 CCGGTTCTGGCAAAAGCTGAGGTGAGTCAAGAGGGTGA-----GGTGAATGTTTG 3752
Db 3954 TCCATTCTCTCAAAAGCTGAGGGGACTCAAGAGGCTGACAGATATGTGATGAAGAAACC 4013
Qy 3753 GATGAGAAAAAGTCAAAAGAGACAGAGGTGTTTGAACACTTGG-----ACCCAAC 3806
Db 4014 AAAAGAGTCACTTTTTCAGAGACTTGAAGGGGTCTATGACACAGGCAATTAACAGTCAAT 4073
Qy 3807 AGTCAAAAGGCTGCTGATGTGCAATGACAGTGAAGTATGGAGTGGCCGGGTGTCAAG 3866

Db 4074 CGGAAAAAGTCACTGAAGTTGCCCTTAAAGTGAAGGACAGAGAAAGCTGAATGTAA 4133
Qy 3867 GAAAAAGAGACTGTAAGTGCAGAG-----TTTAACTTGGAGAGGGA 3911
Db 4134 AAGATATGCTCTTGAACCTGACAGAGTCAAGTAACTCTCCATCCATCCCGGTGAGAGA 4193
Qy 3912 GAGATGGAACCTGACGTTGAAAGAGAGAAAGGAGACAAAGCCAGAGCAAGTATGTAA 3971
Db 4194 GAGATGTGATTCAAGTCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4253
Qy 3972 GAAAG---TGAGCAGAGAAACAGCCGCTCTGAGCATGAAGAGACCTACGGGAAGCCAGTC 4028
Db 4254 GAGAAAGCTTGAAGACAGAAACAGCTGTACGTAATCTGAAGAGTCAATGAAGACACTCTC 4313
Qy 4029 CTGAACCTTGACATGCCCAGCTCAGAGAGGGGAGAGCAGCTGGAGACCTTGAAGAGAGC 4088
Db 4314 CAGACAGTGAATGTCCCATCATATGAGGGCAAGAGAGAGTCAAGTGTGGAAGAGAGC 4373
Qy 4089 CCTTCTC---TCCGAGACCAAGCAAAAGCAGGTTCATGAGGTTCAAGTTCAAAGCTG 4145
Db 4374 CTTCTCTCTGCTGAGTCAAG 4433
Qy 4146 GACACAAACAGTCACTCAAAACAGCAGAAAGCTGTGAGAAA-----GTCATAGAAACGATT 4199
Db 4434 GAGGCAATCATTTCACTTAACAGCGGCTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4493
Qy 4200 GTGATTTCAAGACAGGTGAAGTTCAGAGTGTGTAGGTGCACTTATTTACAGCTGAG 4259
Db 4494 AACATTTTGAAGAAACAGGTGAACCTTGAGAGCTCAGGTGACATTTATGTTCTGGAAGAG 4553
Qy 4260 AAGTCTCTGGAAGAGGGGGGCACTGAGACTTTTCAAGATGACAGAGAGAGAGAGAGAGAG 4319
Db 4554 AATCTCTGAAAAAATGAAGAGTCTTGGCGCTCATCCAGGGAGAGAGTGTGTGCCACA 4613
Qy 4320 GGGCTGAGTCTCAGAGAGAGATCCATCCATATGATTAATCTCTGCTCTGAAAGAGACC 4379
Db 4614 GGGCCCGACCTGTCAAGGAGAAATGCACACAGTATGATTTGTCTATCCAAAGAGAGC 4673
Qy 4380 CTACATCTGACTTCAAGAGAGAAATTAAGCGCATCCAGAGAGAGAGATCAGAGAGAG 4439
Db 4674 TTAAGTTCGAGCTGGAAGGAGAGAGAAACCATCATCAGAAAGTGAAGTCAAGTGAAGTC 4733
Qy 4440 GACAAAGCAGATGCTGTGCTGATGCTGACGGCAAGAGAGATTAACAGCAATGGAAGAGTC 4499
Db 4734 GATGAGCAGGTGCTTGCC-----AGAGGTCAAAAGTATGATGAATG---AGAT 4784
Qy 4500 CTCAAGCTGAACCTGAGATCTCTGAACCTTGAAGTAAAGAGCAACAAGATTGTGCTGAC 4559
Db 4785 TTAGAGCTTGAAGATTTGGATTTTGAACCTTGAACCAAAAGCAGTAACTGTCCAAAGC 4844
Qy 4560 GTCAITTCAGACAGCCGTTGACCAAGTTCGACGTAC---AGAAACAGCCCGGAAACTCAT 4616
Db 4845 ATCATTCAGACAGCCGTTGACCAAGTGTGATGATCAGAGAGAGAGAGAGAGAGAGAGAG 4904
Qy 4617 GCTTATGATTCACAGCCAGCTTC-----CTGCATGACAGCTTGAAG 4660
Db 4905 ACGTCTGATTACAGACCAAGCTTCAGTGAATAAGCTGACACCCAGAGAGCTGAGAGAG 4964
Qy 4661 CAGGAGGCCCAACAGATGCTGAGCAAAATAAGATGCCAAGATGGAACCAAGAGTGC 4720
Db 4965 GAAACGAGAAAG 5024
Qy 4721 GCAAGCCA-----GAGAGACTTGTCAAGTCTGACGTTCTGGA 4759
Db 5025 TCAGCCAAAGAGAGATCAGAGTCAACGCAAGTGGAGACAGACATTTGATATTTCCAA 5084
Qy 4760 GGCATGGGCTCAGCTCGGAAATGCTTCCGCGCTTGTGAGTGAAGGCGCGGTGCA- 4818
Db 5085 GACATGATGAAGCTCAGAAAAAGACATGACTGTGAGTGAAGAGGTTTCACTGTAAAT 5144
Qy 4819 ----- 4818
Db 5145 GATCAGCAGCTGGAAGAGAGTGTCTCTCCATCTGAGAGAGAGAGAGTGAAGCTGGAACA 5204

Qy	4819	-----AGTAAGCATTTGAGAAGCTGCCT	4840
Db	5205	AAGTCTGTGCGAAGAAGATGATGTCATGCTTTGTAGCAGAAAGAAATAGAGAAGTCACTA	5264
Qy	4841	CCTCAACCCCAAAGA--TCCAAAAGGAGCATGCTGTGTATGGCCCTCAGCTCCAAAGCTTA	4898
Db	5265	GTTGAACCGAAAGAAGATGAANAAGGTGATGTTGTATGACCTGAAACCCAGAACTCA	5324
Qy	4899	GCCAGGACAGGCGAGTGCCTCTGGAACCTTAACAAGAAGATCCCCAGACACCAACGGA	4958
Db	5325	GCCCTGGCTGATATGATGCTCAGGAGGCTTTAAACCAAGAGTCCCCAGATACAAATGGA	5384
Qy	4959	CCTAAGCTTAACCCAGAGGAGCGATCCCCAAAGTTGAGGTCCTCAGGAGAGAAATGCT	5018
Db	5385	CCAAAAAAGAGAGAGGAGTCCCAAGAGTAGAATTCGAGGAAGGAAAAAGTGCAC	5444
Qy	5019	ACCAAGTCAGTCAAAGAGAACAGGCGCCAGGAGAGAGGAGCTGCGAGGAGCCAAAGGGA	5078
Db	5445	AGTGAATTCAGATAAAGCGATCAACACCCCAAGCACAGGAGGAGTTACAGAAACAAGAGAGA	5504
Qy	5079	GACCTGCGCAGAACTCTAAGATGTTAGTG-----CTCATTTGTAC	5117
Db	5505	GAATCTGCAAGTCAGAACTTACAGAACTTTAAAAACATCATGCAAGTTAAACTCATTTGCT	5564
Qy	5118	ATCTGTAAGACCAAGATGTGAAACCAAGTCACAGAACAA---GATGCTGCTGTTGGAC	5173
Db	5565	GTTTGGAGACCAAGATGTGAAGCAAGTAGTAGAAGAAATGAATGCTGCTGTGAGA--	5623
Qy	5174	CTTGAGACCAAGATTTTCAGAGGCCATGAGATCCAGAGAGCGGCGCTCCAAATGATTTCC	5233
Db	5624	CTGAAGACCAAGTATTTTCAGA-ACCTTTGAGAAATTGGAGAGCGGCACATCACTGATCTCA	5682
Qy	5234	ACCCAGTAGAGCACCCGCGAATTCGAGGCTTCATTCGGAGCTAGAGCCAGCTTAACATT	5293
Db	5683	TTTCTAGAGAGC--CCCTGACAACTCTCGAGGCTTCATCAGAGCTAGAGCCATTTAACTT	5741
Qy	5294	TCCTCGTTTCAAGACTGCCTTTGATTTGCGCCCTTCATGCGCGTCCGTGATTTCTTAACATT	5353
Db	5742	TCCTCTTTCAAGACCAAACCTACATTTTCCCTTGATTAAC-----	5781
Qy	5354	TCCTCGTTTCAAGACTGCCTTTGATTTGCGCCCTTGATGCGCGTGGTGTATTTTCGGATTTA	5413
Db	5782	-----CATATAAATTTCTGATTTA	5799
Qy	5414	AGGTCTGCGCTTCTCAACCTGGNAC--CAATTCGCGCATACCTAGTTTCCATTTCTCAAACT	5472
Db	5800	AGGTCTTAAATTTCTTAACTCGGAATGGAGTTGGCAATACCTAGTTCTGCTCTCGAACT	5859
Qy	5473	GGAGCATCTCCTTTATGTATTTATATGATGATGTTTTATGTAGTCTCTCTCTGTAACCTAT	5532
Db	5860	GGAGTATCATCTTTACATATTTATATGATGATGTTTTAAGTAG---TCCTCTGTATCTAT	5916
Qy	5533	TGTATATTTTTTTTCTAACGTTT-----AAGCACATGCTTTTTTGTATTA	5575
Db	5917	TGTATATTTTTTTTCTTAATGTTTAAAGAAATGCGAGGATACTACATGCTTTTGTATCA	5976
Qy	5576	TGCAATATATAACGGGTGTCAGCCATACGAGCGCTTTGAAAAGTCCAGGCTCAAACTG	5635
Db	5977	CACAGTATATGATGGGCGATGTGCCATAGTGCAGGCTTTGGGAGCTTTAAGCCCTCAGTTA	6036
Qy	5636	TAACTGCGACAAACAGATAAACATTC-----CTGGCAAGAGAGACAGACTTTTTT	5686
Db	6037	TATAACCCACAAAAACAGAGCTCTAGATGTAACTTCCTGATCAAGGTACAAATTCCT	6096
Qy	5687	TAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCCCTCTGAAAGTGGTGTGTTTTCCCT	5746
Db	6097	TAAAAATTCATAATGATGAGGTCCATATTTAGTGGTACTCTGAAATTTGGTCACTTTCTCT	6156
Qy	5747	ATGCACAGGAGCTCAGAAATAAAAAACCCATTTTGAACATCATCAGAGATGCCCAATATT	5806
Db	6157	ATTACGCGAGTGTGCGAAACATTAAGAGCATTTTGAACATACAGAAATGTTCTATTGTC	6216

QY	5807	ACCATGATTTTTTCCCCCTTTTTTGGCTTATCCAGTCCAGGTTGGAAAGAGTCTCCTCTG	5866
DB	6217	ATTGGGAAATTTTGC-----TTTCTAACCCAGTGGAGGTTAGAAAAGAGTTATATTCT	6269
QY	5867	TGTCAGA-----TTAAGCCCTCTCTTAAATGATATGGAACAATGAGT	5909
DB	6270	GGTAGCAAAATTAACCTTTACATCTTTTTTCCCTACTTGTATGGTTGTTGGACCGATAGT	6329
QY	5910	GTGCTTAAGGCCATGAGA-----TGTTTCTTAATGCAGAGGAATCTGT	5953
DB	6330	GTGCTTAATCTGAGGCAAGTAGTGAATATGTTTATATGTTATGAAGAAAGAAATGT	6399
QY	5954	TGTAGCTTTTTTGTATGTTACTCTTATGCTGACCGAATTCATATGCAGATCGAAGTG	6013
DB	6390	TGTAAG--TTTTTGATTCTACTCTATATGCTGAGTGCATTACACATGGCATGAAATA	6447
QY	6014	AGTCTGTTCTTTTACAGATGCTATTTTATAGATACATGGAGTTTGTCTGTTTATATCTG	6073
DB	6418	AGTAGGTTCTTTACAAATGGTATTTTATAGATACTGGATTGTTTGTGCCATATTTG	6507
QY	6074	TGCCCTTTCTTTAAGAACAAATGTTGCATTTATGTTCTTTTGGATAAAATGTTGATTGACAA	6133
DB	6508	TGCCATTCCTTTAAGAACAAATGTTGCAACATTCATTTGGATAAGTTGTGATTTGACGA	6567
QY	6134	CTGATTTAAATAAA 6147	
DB	6568	CTGATTTAAATAAA 6581	
RESULT 3			
US-09-724-676-12735			
; Sequence 12735, Application US/09724676			
; GENERAL INFORMATION:			
; APPLICANT: CompuGen LTD			
; TITLE OF INVENTION: Variants of alternative splicing			
; FILE REFERENCE: 129181.4 CompuGen			
; CURRENT APPLICATION NUMBER: US/09724,676			
; CURRENT FILING DATE: 2000-11-28			
; NUMBER OF SEQ ID NOS: 97222			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 12735			
; LENGTH: 5931			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-724-676-12735			
Query Match		32.4%;	Score 1994.8; DB 5; Length 5931;
Best Local Similarity		67.6%;	Pred. No. 0;
Matches 3281;		Conservative	0; Mismatches 1332; Indels 237; Gaps 25;
QY	3	GGCCGGCGGGAGTAGAGAGCCACTGAGCCATGGCGCAGGACAGTTCACCGAGCAGCGG	62
DB	162	GGCTAGGCGCGGAGAAAGTGCGGAGAGCCATGCGGCCCGGAGACTCACCGAGCAGCGC	221
QY	63	AGCCCGCAGCAGCCG---GCGGGGAGCGACACGCGCAGCAGCTGTGCTCAGTGGCCAT	119
DB	222	AGCCCGGAGCAGCCGCCGAGGGAGCTCCACGCGGCTGAGCCCGAGCCAGCGCGCGC	281
QY	120	GGGCCCGCAGCTGAAGCCTCGGGA-----GCAGCTGGAGACCCCGCCGACCGCGAC	170
DB	282	GGCCCTCTGGCGGAGGGCGGCCACAGACACACCGCGAGCCCGCCATCGTCTCGGAC	341
QY	171	CCGCCCAAGCTCCACAGAAAGTGGCCAGCTGCTTCTGTCAACGGCTAGCTGAA	230
DB	342	CCGCCCAAGCTCTTACAGAAAGTGGTCAGCTGTCCACCATCAATGGCGTAGCTGAG	401
QY	231	CAAGGAGATGTCATGTCCAAAGAGGAAAAACGAGGG-----	268
DB	402	CAAGTAGCTCAGCCTCCAGGAGGGTGACCTAAATGCGCCAGAAAGAGCCCTGAACGGT	461
QY	269	-----GCAGAGGAAGAGTCTGTTGATGAGGATGTTGGACAGCA	308
DB	462	CAAGGAGCCCTAAACAGCCAGGAGAGAAAGTCAATTGTCAAGAGGTTGGACAGCA	521

QY 309 GAGTCAGAAATGTGAGAGAAAAAGCCGAGTTTGAAGAAATGCGGCCAATCTCCACGCT 368
Db 522 GACTCTGAAGATGTGAGGAAAGAGATCCGATTAAGATGCGCTACTAAGTCAACGCGTT 581
QY 369 GTTGAAGATATCAAAAGATGGGAGAGAGGAGCATCAGAAATTAATTGAACAGATCCCT 428
Db 582 GTTCACGACATCAAGATGATGCGAGAGAGACACCCGAAATTAATGACACAGATTCTT 641
QY 429 GCTTCAGAAAACAATGTGAGAAATGATACAGCCTGTGATCCAGGCTAATGATGTT 488
Db 642 TCTTCAGAAAGCAATTTGAAGAGCTTAACACAAACCACTGAGTCCAGGCTAATGATTT 701
QY 489 GCGTTCAAGAAAGTATTTAAATTTGTGTTTAAATTCACGATGAAGAGATTAAT 548
Db 702 GGATTTAAGAAAGTGTTTAAGTTTGTGCTTTAAATTCAGTGAATAAAGATTAAGCA 761
QY 549 GAAAAGTCAAGTACTGTCCAACTCACTGTCAAGAAAGATGAAGGCGAAGGCGCAGAA 608
Db 762 GAGAAAGCTTGACACTGTCCAGTACTCACTGTGAAGAAAGATGAAGGCGAGAGCAGCA 821
QY 609 GCTTCTGTGAGAGCTGAGACCAACGAGGCCAGTGTGAGACTGCGTCCGAGAGTCA 668
Db 822 -----GGGGCTGGGAGCCACAGAGCCCAAGCTTGGGGCT-----GGAAGAGCA 866
QY 669 GCATTCAAAAGAAATGTAGCTGAAGCAATTCACAGAGAACAAAGGACCTCTGAAGCAA 728
Db 867 GCATTCAAAAGAAAGCGAAACCCAAACATCTACAGAGAAACCGAAGAGACCTCTGAAGCCT 926
QY 729 GAACGAGCAGACAGAAATCCCTTCAGAGCGAATCTGATCAAGCGGCTGAGAGAA 788
Db 927 GAGCAAGCCAGCAGAAATTTCTCCCAAGCGAATCTGGCGAA---GCAGTGAGAGAA 983
QY 789 GCCAAAGTGAAGAGAAAGAAACAAAGAAAGAGCCCAAGTCCCAAGATCCCG 848
Db 984 TGCAAAAGAGAAAGAGAGAGAAACAAAGAAAGAACTGACAGTCTGACAGATCTCCG 1043
QY 849 AGCAGCCAGTCAACAGTGAAGCAATCTTCTTCAGAAAGTCTTCACTACGCTTGG 908
Db 1044 ACTATCTCCGTGACCAAGTGAACAGATTCACCTTCAAAATTTTTCATCTAAAGTGG 1103
QY 909 GCGGCTGGCGCAAGAGCCAGCTTCAGAAATCAAAAGAGATGATCTGAAACTGCC 968
Db 1104 GCGGCTGGCGCAAAAGACCAATTTTCAGAGAGCCGAAAGAGATGAACTGAACTTCA 1163
QY 969 GAGAAAGAAAG 1028
Db 1164 GAGAAAGAAAG 1223
QY 1029 CCAGCTTCGAGAG-----GAGAGAGCCGCGAGAGAGC 1061
Db 1224 GTTGCTTCGAGAACTGACCGCTCCGAGCAAGCCCAACCAAGAGCCGCGCAAAAGT 1283
QY 1062 ACAGCCAGGCGCAGGTGTGACGACACTACGAGAGGTGAGAGCTGCTTGGAAAGCAG 1121
Db 1284 GCCCAGAGCCCGGTATACAGCTGAATGTGAAGAGTTGAGCTGCTTCAGAGAGAGCA 1343
QY 1122 GTTGTGACCTGAGAGCAATCTGACAGAGAGAAAGTGTCTCTTTGGCAACGAGAGTGT 1181
Db 1344 GTCAATGCTGCGACAGGACCTTCTGAAGAGAAACCTGCTCGTTGGCGACAAAGTGT 1403
QY 1182 GATGAGAAAGTGAAGCCCAAGAGAA---GTTGTTGACAGAGTCAAGTGAAGCAGTGG 1238
Db 1404 GATGAGAAATTAAGAGTCCACCAAGAGAGAGTGTGTGCGCGAAGTCCAGTCAAGCAG 1463
QY 1239 GAG 1298
Db 1464 GAG 1499
QY 1299 GGAACAGAGAGATCTTGGCCCTGAGAAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 1358
Db 1500 GAAACAGAGAGAGTCTGTGACAGCTGAAGATTTGTTGAATGGAATGCAAGAGAGAG 1559

QY 1359 GCTGAGCCTGTGAGAGAGCTGATGAAGAGAGAGATGTGTCTCTGAGAGAGACAC 1418
Db 1560 GCCAACCTGCGCAAGAGAGCTGTGAAGCTCAAGAAACGTGTGTTTCCGAGAGAGACCT 1619
QY 1419 ACTCAACTGACAGACCTAAGTCTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1478
Db 1620 ACACAGAGAGAGCTGACCTCAAGTCTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1679
QY 1479 GTCAGTGAAGTGAAGATGTCTCTCTGAGAAAGATCAAGAGTGAAGAGAGTCCCTG 1538
Db 1680 GTGAGTGAAGTGAAGATGTCTCTCATCAAGAGAGAGATGAAGAGTGAAGAGAGAGAG 1739
QY 1539 AAGAACTCTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
Db 1740 AAGAACTTTTACAG 1799
QY 1599 CGAGAGAGTGGGAG 1658
Db 1800 AGAGAG---GGAG 1856
QY 1659 GAGAGTGTGATGAG 1718
Db 1857 GACAGCAG 1916
QY 1719 ACCAGTGTCTGAG 1778
Db 1917 ATCAAGTGTCTGAG 1976
QY 1779 ACTACTTCGATGAG 1838
Db 1977 GCTACTTCGATGAG 2036
QY 1839 ATGCTGACCCCAAG 1898
Db 2037 ATGCTGACCCCAAG 2096
QY 1899 GAGAGAGTCAAG 1958
Db 2097 GACAAAGTCAAG 2156
QY 1959 GAAGTCAAACTGTTGAG 2018
Db 2157 GAATGAAG 2216
QY 2019 TCAAGTCTTGGAG 2078
Db 2217 TCAGATCTTGGAG 2276
QY 2079 TCCCTTCAGATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2138
Db 2277 TCCCTTCAGATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2236
QY 2139 GAGGACAG 2198
Db 2337 GAGGACAG 2336
QY 2199 GACCAAGCGCAG 2258
Db 2397 GATCCAGGCGCAG 2456
QY 2259 GTCTTCACCTTGGAG 2318
Db 2457 GTTTCACCTTGGAG 2516
QY 2319 CTGGAAG 2366
Db 2517 CTGGAAG 2576
QY 2367 ATCAAGCAGATGAAG 2426
Db 2577 ACTGAACCCGAG 2636
QY 2427 AAGAAAGGCGAG 2486

Db 2637 AAGAAAGGCGAGATGGGAAACAGAAACAGCCCTGTTGAAGCGGCGGCAACAGGG 2696
 Qy 2487 ATAAATGAGGACGACCTTAATGTCCTCCAGCGCTGCTGCTCTGCTGAGTAAATGCAAGT 2546
 Db 2697 GCCAACGAAGATGACTCTGATGTCCTCGGCGCTGCTGCTCTGCTGATGATGCTGTA 2756
 Qy 2547 GAGAGGGAAGATGG-----AAGCCAGGGGAATACGAGCTGCCCAAGCTGCTG 2597
 Db 2757 GAAAGGGAGAAAATGGAGGACACAGCAGCCCAAAAAGCGCAGAGCAGCCGAGCAGAAG 2816
 Qy 2598 GGGCTGTGTACGTGTCGAGGAGCTCAGTAAGACTCTGCTCCACACTGTGAGTGTGCA 2657
 Db 2817 GCAGCCTGAGGTGTCGAAGAGCTCAGCGAGTGTGAGTGTATATGATGAGCAGAGCT 2876
 Qy 2658 GTCAATTGATGGGACCGAGGAGCTCACAGTGTGCAAGAGCGGTCTCCTCTGCTGATATCC 2717
 Db 2877 GTCGTCAGCGGACGAGGCGAGCTACCATTTATTGAAGAAAGGTCTCTCTCTGATATCT 2936
 Qy 2718 GCTTCGTAACAGAACTCTTTGAACACACAGCGGGAGAACCCATGCCACCTGTTGAAGAG 2777
 Db 2937 GCTTCAGTGAAGAACTCTTTGAACAGTGAAGCTGAAGCGGCACTGTTAACTGAGGAG 2996
 Qy 2778 GTCACTGAAAAGACATCATTTGC---AGAAGAACTCTGCTCACCACAGAGTTACCA 2834
 Db 2997 GTATTGGAAGAGAAGTAATTGAGAAAGAAACCCCCACGGTTACTGAACTCTGCCA 3056
 Qy 2835 GAGGTTAAGATGCCATGACGACATGGTCACCAGTGAAGTGAATTTCACTCAGAAAGCT 2894
 Db 3057 GAGAACAGAGAGCGCGGCGGACACGCTGTTAGTGAGCGGANTTACCCCGAAGCT 3116
 Qy 2895 GTGACAGCCACAGAGACTCTCAGAGGCTCTCCGTACTGAAGAAAGTTACCGAAGCATCGGG 2954
 Db 3117 GTGACAGCTCAGAAAATGTCAGAGGCACTTGGGTGCCGAAGAAAGCAACCGAAGCATCTGCT 3176
 Qy 2955 GCCGAGAGACACAGACATGCTGTCGACGTTTCCAGCTGACTGACTCCCCAGACACC 3014
 Db 3177 GCTGAAGAGACCAAGAAATGGTGTGACGAGTCTCCAGTAAACCGACTCCCCAGACACC 3236
 Qy 3015 ACAGAGAAAGCCACCCAGTTTCAGGAGGTAGAGTGGTGTCTAGATACAGAAAGAGAG 3074
 Db 3237 ACAGAGAGGCCACTCCGTTGTCAGAGGTGSAAGTGGCGTACTGACATAGAAGACAA 3296
 Qy 3075 GAGCGCAGACGAGGCGCATCTCCAGCGCTTCGACAGAGGTGAAGAGGAGTCCAG 3134
 Db 3297 GAGAGCGGACTCAAGAGGTCTCCAGGAGTGCAGAAAAAGTGAAGAGGAATCCAG 3356
 Qy 3135 GTCCTCAACC---CAGACTGTGCAGAAACGGGGTCAAAAGCACTGGAGAGAGTTGAG 3191
 Db 3357 CTGCTGGCACCGGTGGCCAGAAAGTGTGCTTACGCTGTGCAGAGAGCAGAGCGAGAA 3416
 Qy 3192 GAGGTAGAGGAGGACTCCGAAGTGTGGCTTCGGAGAAAGAGAGGAGCTTATGCCGAAA 3251
 Db 3417 AGACCAGAGAGCAGGCTGAAGCGTGGGTCTGAGAAAGAGCAGATGTAGTGTGAAA 3476
 Qy 3252 GGACCGGTGAGGAGCTGAGCTGAGCATCTTCGACAGGGCTCTGAGACTGGACAGGCT 3311
 Db 3477 GTAGATGCTCAGGAGGCAAAACTGAGCCCTTTACACAGGGGAAGGTGGTGGGCGAGCC 3536
 Qy 3312 ACTCCAGAGAGCCCTTG---AAGTTCCTGAAGTCAAGCAGCATGTA-----GACCAT 3359
 Db 3537 ACCCCAGAAAGCTTTGAAAAGCTCTCAAGTCAAGAGAGCAGATAGAGTCCAGTGAGCTT 3596
 Qy 3360 GTCGCCAGCTGCCA-----GTTATCAAGCTCCAGCAGCTGATGGAA 3401
 Db 3597 GTAACTACTGTACGCCGAAACCTTAGCTGGGGTAAATCAAGAGATGGTATGGAA 3656
 Qy 3402 CAGCGGTGGCCCTGAGTGCATCCGAACCTTGACAGACAGTGAAGCAAAATGGAAGCACT 3461
 Db 3657 CAGGCTATCCCTGACTCGGTGGAAACCCCTACAGACAGTGAAGTGTGAGAGCACC 3716
 Qy 3462 CCCTTAGCAGATTGACACACTGAGATGGGACACAGCAAGATGAACCATTTGACGCCAG 3521

Db 3717 CCGTAGCCGACTTTGACGCAACAGGCACAACCCAGAAAGACGAGATTTGTGAAATCCAT 3776
 Qy 3522 GACAGTAAAGCCACTGACAGCTGTGAGGAGTCAAGGTCAAGAAAGAGAGCGGCTACT 3581
 Db 3777 GAGGAATGAGGTGCGATCTGTTACCCAGTCAAGGGGACAGAAAGCAGAGCAGTTCTCT 3836
 Qy 3582 GCTCAGAAAGAGGAGGCTTGCACACTACCTAATATGTTTCCAGCCAGGAACATGGG 3641
 Db 3837 GCACAGAAAGAGGAGGCTCCAGCACCTTCAGTCTTGTGTTCCAGGAAGAAATAAGAA 3896
 Qy 3642 GAAGAACACAGGAAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGTCAGCCGTG 3701
 Db 3897 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTCACTGAGAACTGTA 3956
 Qy 3702 CCCGTTCTCGCAAGACTGAGGTGGGTCAAGAGGTGA-----GGTTGACTGGTTG 3752
 Db 3957 TCCATTCTGTCAAAGACTGAGGGACTCAAGAGGCTGACAGTATGCTGATGAGAAAC 4016
 Qy 3753 GATGGAAAAAAGTCAAGAAAGAACAGGAGGTGTTGTACACTCTGG-----ACCCAAC 3806
 Db 4017 AAGACGTATACCTTTTTCGAAGGACTTGAGGGGTCTATAGACACAGGCATACAGTCAGT 4076
 Qy 3807 AGTCAAAAGGCTGCTGATGTGACATATGAGTAAGTGTGGAGTGCAGGCTGTCTAG 3866
 Db 4077 CGGAAAAAGTCACTGAGTGTGCCCTTAAAGTGAAGGACAGAAAGCTGATGTAATAA 4136
 Qy 3867 GAAAAGGAGAGTACTGAGTGCAGAG-----TCTTAGCCTGGAGGAGGGA 3911
 Db 4137 AAGGATGATGCTTGAACCTGACAGTCAAGCTACGCTCTCTCCATCCCTCGTGGAGAGA 4196
 Qy 3912 GAGATGGAATGACCGTTGAAAAGGAGAAAGGAGACAAAGCAGAGCAAGTGAAGTGA 3971
 Db 4197 GAGATGTTAGTTCAAGTCAAGGAGGAGAAACAGAAAGCAGAGCCACCTGTGAATGAA 4256
 Qy 3972 G---AAGGTGACGAGAAACAGCGCTCTGAGCATGAAGGAACCTACGGGAGCCAGTC 4028
 Db 4257 GAGAGCTTGGAGCACGAAACAGCTGTTACCGTATCTGAAAGGTCTAGTAAGCAGCTCTC 4316
 Qy 4029 CTGACACTTGAATGCCAGCTCAGAGAGGGGAGGCACTGGGAAGCCTTGGAGGAAGC 4088
 Db 4317 CAGACAGTGAATGTGCCCATCATAGATGGGCAAGGAAGTCAAGCAGTTTGGNAGAGC 4376
 Qy 4089 CTTTCTC---TCCAGACCAAGACAAAGCAGGTTGCATAGAGTTCAAGTTCAAAGCCTG 4145
 Db 4377 CTTCTCTCTGCTGCTAGGTCAAGAGGAGGAGTATGCACCAAAATTCAGTTCAGAGCTCT 4436
 Qy 4146 GACACACAGTCACTCAACACAGCAGAGCTGTGGAAGAGTCTAT-----AGAAACGGTT 4199
 Db 4437 GAGGCATCATTTCACTTAAACAGCGCTGCAGAGGAGGAAAGGTCTTAGGAGAAACTGCC 4496
 Qy 4200 GTGATTTCAAGACAGCTGAAAAGTCCAGAGTGTGAGGTGCACACTTATTACCAGCTGAG 4259
 Db 4497 AACATTTTGAAGACAGTGAACGTTGGAGCCTGAGGTGCACATTTAGTTCTGGAAGAG 4556
 Qy 4260 AAGTCTCTGCAACCGGTTGGCCACTGGACTTTTCAGCATGCAAGAGACACGTPACCCCTG 4319
 Db 4557 AAATCTCTCTGAAAAAATGAAGACTTTGCGGCTCATCCAGGGGAAGATGCTGTGCCACA 4616
 Qy 4320 GGGCTGAGTCTCAGGCAGAAATCCATCCCAATCATATGATTAATCTCTCTCTGAAAGCACC 4379
 Db 4617 GGGCCGAGTGTGAGGCAAAATCGACACCAAGTGTATGATCTGCTACTACCAAGAAAGGC 4676
 Qy 4380 CTACATCTCTGACTCAAGGAGAAATAAGCGATCCAGAGAGCGGATCAGAGGAAGAG 4439
 Db 4677 TTAAGTTCGACCTGGAAGAGAGAAACCAATCACTGAGTGGAGTCAAGTGAAGTC 4736
 Qy 4440 GACAGCCAGATGCTGGTCTTGTGATGCTGACGGCAAGGAGAGTACAGCAATCGAAAAAGTC 4499
 Db 4737 GATGACAGGTGCTTTCG---AGGAGTCAAAAGTCAAGTGTAGCAATTTGAGGAGGAT 4790
 Qy 4500 CTCAGGCTGAACCTGAGATCTCTGGAATCTTGAAGTGAAGCAACAAAGATTGTGCTGAAC 4559
 Db 4791 TTAGAGCTGAAAAATGGGATTTTGGAACTTGGACCCAAAAGCAGTAAACTTGTCTCAAAAC 4850

Oy	4560	GTCAATTGAGACAGCCGTTGACAGTTGCGACAGTAC	---	AGAAAACAGCCCCCAAACTCAT	4616
Oy	4561	ATCATCCAGACAGCCGTTGACAGTTGTCATCGTACAGAAAGAAACACCCCGAAATGTTG			4910
Oy	4617	GCTATATATTACAGACCCAGAGTTCTTCGATGACGAGGCTGACAGCAGGGA			4666
Oy	4911	ACGCTGTAGTTACAGACACCAAGCTCAGCTATATATAAAACGTGACAGCCAGGA			4960

RESULT 4

```

US-09-724-676A-12735
; Sequence 12735, Application US/09724676A
;
; GENERAL INFORMATION:
;
; APPLICANT: Compugen LTD
;
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12735
;
; LENGTH: 5931
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
US-09-724-676A-12735

```

Query Match	32.4%	Score 1994.8	DB 5	length 5931
Best Local Similarity	67.6%	Pred. No. 0		
Matches 3281; Conservative	0	Mismatches 1332	Indels 237	Gaps 25

Qy	3	GGCCGGCGGGAGTGAAGAGCCATGAGCCATGAGCGCAGAGCAGTTCACACGAGACGCG	62
Db	162	GGCTAGGGCGGGAGGAGTGGCGAGGAGGCATGGGCGCGGAGGCTCCACCGAGCAGCGC	221
Qy	63	AGCCCCGAGCAGCCG--GCCGGAGCGACACCGCAGCGAGCTGGTCTCAAGTGCCAT	119
Db	222	AGCCCCGAGCAGCCGCGCGGAGGGGAGCTCCACCGCGGTAGGCCGAGGCCAGCGCGCGC	281
Qy	120	GGGCGCCGAGCTGAAGCTCGGGG--GCAGCTGAGAACCCCGCCACGCGGAC	170
Db	282	GGCCCTTGGGCGGAGCGGCGCCGACACACCGCGGACCCCGCCATCCGTGGCTTGAGAC	341
Qy	171	CCCGCCACCAAGCTTCCACAGAGAATGCGCAGCTGTCTTGTCAACGGCTGATGAA	230
Db	342	CCCGCCACCAAGCTCTTACAGAGAAATGGTCAGCTGTCCACATCAATGCGGTAGCTAG	401
Qy	231	CAAGGAGATGTCATGTCACAGAGAAACCGAGGG-----	268
Db	402	CAAGATGAGCTGAGCTTCAGAGAGGGTGACTTAATGGCCAGAAAGAGCCTGAACGCT	461
Qy	269	-----GCAGAGAGAAAGACTCGTTGATAGAGATGTTGGACCGCA	308
Db	462	CAAGGAGCCCTAAACAGCCAGGAGAGAAAGAAAGTCAATGTTCAACAAGTTGAGCAGAGA	521
Qy	309	GAGTCAGAAAGATGTGAGGAGAAAAGACCGAGTTGAAAGAAATGGCGGCCAATCCACAGT	368
Db	522	GACTGTGAAGATGTGAGGAAAGAGACTCCGATTAAGAGATGCTCTAAGTGAAGCTT	581
Qy	369	GTTGAAGATATCACAAAGAGATGGCAGAGAGAGACATCAAAAATTAATTGACAGATCCCT	428
Db	582	GTTTCACGACATCACAGATGATGGCGACGAGGAGACACCGAAATTAATCGAACGATCTCT	641
Qy	429	GCTTCAGAAAACATGTGGAAGAAATGGTACAGCTCTGTAATGCCAGGCTAATATGTT	488
Db	642	TCTTCAGAAAGCAATTTGAAGAGCTAACCAACCACTAAGTCCAGGCTAATGATAT	701
Qy	489	GGCTTCAGAAAGATTAATAATTTGGTGTTAATTAATTCACGGTGAAGAGATAAAT	548
Db	702	GGATTTAAGAGGTGTTTAAGTTGTTGGCTTTAAATTCACTGTGAAGAAAGATTAAGACA	761
Qy	549	GAAAAGTCAGATACTGTCCAACTCACTGTCAAGAGAGTGAAGGCGAGGGGTCAGAA	608

Qy	1659	GAGAGTCTGATGAGCAGAGAGGCGAGAGAGCTCTGGCTCGTCCCGAGGAGCGCTGAGGAG	1711
Db	1857	GACAGCCAGGAGGAGCAAAAGGCGGAGAGCTCTGGCTCATCCCTTGAGGAGCCCGAGGAG	1916
Qy	1719	ACCAAGTGTCTGGAGAAAGGGCGCTGGGAAGCACCCAGGATGGGGAAGCTTGAGGAAGGA	1778
Db	1917	ATCAGGTGTCTGGAAGAAGGGCTTACCGGAGTGCACGAGGATGGGAAGCTGAGAGAGGA	1976
Qy	1779	ACTACTTCGATGGAGAAAGAGAGAGAGAGGATCATCTCTCTGGGCATCTCTTCAAAAG	1838
Db	1977	GCTACTTCGATGGAGAAAAAGAGAGAGGTGTCACTCTCTCTGGGCATCATTTCAAAAG	2036
Qy	1839	ATGTGTACACCCAAAGAAACGGGTCCCAAGACCTTCTGAGAGTGAACAAGGAGAGAGCTG	1998
Db	2037	ATGTGTAGGCCCAAGAAAGCGTGTGACGGCTTTCGGAAGTGTAAAGAAAGATGAGCTG	2096
Qy	1899	GAGAAGGTCAAGAGCGCCACTTTGTCTCTCCACTGTATAGCACAGTGTCAAGAAATCCAAGAT	1958
Db	2097	GACAAGGTCAAGAGCGCTACTTTGTCTTCCACCGAGAGCAGCTCTTGAATTCAGAA	2156
Qy	1959	GAAGTCAAAACTGTTGTGAGGAACAAAAGCCAGAGGAACAAAGCGTGGGTGGATACT	2018
Db	2157	GAATGAAAGGAGCGGTGGAAGAGCAAAAGCCGGAAGAACAAAGCGCAAGTGGATACC	2216
Qy	2019	TCAGTGTCTTGGGAAGCACTGATTTCTGTCTCGATCATCCAGAAGAGAGCAAGGAAGCA	2078
Db	2217	TCAGTATCTTGGGAAGCTTTAATTTGTGTGGGATCATCCAGAAGAGCAAGGAGAGGG	2276
Qy	2079	TCCTCTTCAGATGATGAAGGAGGCGCCAAAGCACTCGGAGGGGACAGTGCACAGAGCAGAG	2138
Db	2277	TCCTCTTCAGTGAAGGAGGGGACCAAAGCAATGGGAGGAGACCACAGAAAGCTGAT	2336
Qy	2139	GAGGCCAGAAAGACAAAGAACCGGAAACAGACGCTGTTCTGCCAGCACCCAGAGGAGAG	2198
Db	2337	GAGCCGCGAAAGACAAAGAGACGGGACAGACGGGATCTCTTGTGTCTCCCAAGAACAT	2396
Qy	2199	GACCAAGCGCAAGGAGTTCCTCACCGAGCAGCGGGAAGCCCTTCCGAAGGGGAGAGT	2258
Db	2397	GATCCAGGCGAGGAAGTTCCTCCCGAGCAAGCTGGAAGCCCTTACGAAGGGGAGGGC	2456
Qy	2259	GTCTCCACTTGGGAGTCATTTAAAGATTAGTCACCTCCAAGAAAAAAATCCAAGTCAAAA	2318
Db	2457	GTTTCCACCTGGGAGTCATTTAAAGGTTAGTCACGCAAGAAAAAATCAAAGTCCAG	2516
Qy	2319	CTGGAAGAGAAACCGAGACCTCCTAGCTGGGTCTGGGTAGAACATTCATCTCCAGAC	2366
Db	2517	ATCGAACCGAGTAGAAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGAGCGCGG	2426
Qy	2367	ACTGACCCGGTAAGAGAAATCCTGGGTCTCAATCAAGAGTTTATTCCTGGAGCAAGG	2486
Db	2577	AAGAAAAAGGCGAGACGGGAAGCAAGGACCACTGTGGAGACTCAGGGCCAGTGGAG	2636
Qy	2427	AAGAAAAAGGCGAGATGGAAAAAGAAACAAGCCCTGTTGAAGAGCGAGGSCCAACAGG	2486
Db	2637	ATAAATGAGACGACCTAATGTCCAGCGCTGTGCTCTGTCTGAGTATATATGCAAGT	2696
Qy	2487	GCCAAACGAAGATGACTGTGATGTCCCGCGCGTGGTCCCTCTGTCTGAGTATGATGCTGA	2546
Db	2697	GAGAGGAGAAAGATGG-----AAGCCCAAGGGGAATACGGAGCTGCCCCAGCTCTG	2756
Qy	2547	GAAGGGAGAAATGGAGGCACAGCAGGCCCAAAAGGCGAGAGCAGCCCGAGCAGAG	2597
Db	2757	GGGCGTGTGATCGTGTCCGAGGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTGCGCA	2816
Qy	2598	GCAGCCACTGAGGTGTCCAAGGAGCTCAGCGAGAGTTCAGTTTATATGATGGCAGCAGCT	2657
Db	2817	GTCAATTCATGGGACCGGCGAGTCAACAGTGTGGAAGAGCGGTCTCTCTGTGATATACC	2876
Qy	2658	GTGCTCAGCGGACGAGGGCAGCTACCATTTATGAAGAAAGGTCTCTCTTCTGATATCT	2717
Db	2877		2936

Qy	2718	GC	TTCCGTTACAGAA	CCTCTTGACACACAGCGGAGAGCCATGCCACCTGTTGAAGAG	2717
Db	2937	GC	TTTCAGTGCAGAA	CCTCTTGAAACAAGTAGAAGCTGAAGCCGCACTGTGTTAACTGAGGAG	2996
Qy	2778	GT	CAC	TGAAAAAGACATCATTCG- --AGAAGAAAC	2834
Db	2997	GT	ATTGGAAGAGAG	AGTAAATTGCGAAGAAGAAACCCCCACGGTTACTGAACCTCTGCCA	3056
Qy	2835	GAG	GGTAAAGATGCC	CATGACGATGTTCCACGAGTGGATTTTCACCTCAGAAGCT	2894
Db	3057	GAGA	ACAGAGAGG	CCCGGCGACACGGTCGTTAGTGAGGCGGAATTGACCCCGAAGCT	3116
Qy	2895	GTG	ACAGCCACAGAG	ACCTCAGAGGCTCTCCGTACTGAAGAAGTTACCGAAGCATCGGG	2954
Db	3117	GTG	CACGCTCGA	GAACATGCGAGGGCCATTGGGTGCGCGAAGAAGCAACCGAAGCATCTGCT	3176
Qy	2955	GCC	GAAGAGACCA	CAGACATGTTGTCGCGAGTTTCCAGCTGACTGACTCCCCAGACACC	3014
Db	3177	GCT	GAAAGAGAC	CAAGAAATGGTGTGACGAGTCTCCCAAGTTAACCGACTCCCCAGACACC	3236
Qy	3015	AC	GAGGAAGCC	ACCCACAGTTCCAGGAGGTAGAGAGTGGTGTGTAGATACAGAAGAAGAG	3074
Db	3237	AC	GAGGAGGCC	ACTCCGTCGAGGAGGTGGAAGTGGCGTACCTGCATAGAGAGCA	3296
Qy	3075	GAG	CGCCAGACG	CAGGCCATCTCTCCMAGCCGTTGCAGACAAGAGTGAAGAGAGGATCCCG	3134
Db	3297	GAG	AGGCGGCACT	CAAGAGGTCTCTCCAGGCAGTGGCGCAAAAAGTGAAGAAGAGAAATCCCG	3356
Qy	3135	GTG	CCTCAACC- -	-CAGACTGTGCAGAGAAACGGGGTCAAAAGCACTGGAGAAAGGTTGAG	3191
Db	3357	GTG	CCTGGCA	CCGGTGGCGAAGATGTGCTTACGCTGTGCAGAGAGCAGAGGCGAGAA	3416
Qy	3192	GAG	TAGAGGAGG	ACTCCGAAGTGTGGCTTCGAGAAAGAGAAAGGACGTTATCCCGAAA	3251
Db	3417	AG	ACCAGAGAG	CAGGCGTGAACGCTCGGTCTGAAGAAGAGAGACGAGTGTAGTGTGAAA	3476
Qy	3252	GG	ACCGGTG	CAGGAAGCTGAGCACTTTGCACAGGGCTCTGAGACTTGAAGCAGGCT	3311
Db	3477	GT	AGTGTCTCAG	GAGGCAAAAACCTGAGCCCTTTTACAAGGGAAGGTGGTGGGCGACACC	3536
Qy	3312	ACT	CCAGAGAGC	CTTG- --AGTTCTGGAAGTCAAGGCGAGATGTA- -----GACCAT	3359
Db	3537	ACCC	CAAGAAAGCTTT	GAAAAGCTCTCAAGTCAAGAGAGCATAGAGTCCAGTGAGCTT	3596
Qy	3360	GT	CCACAGT	GGCCA- -----GGTTATCAAGCTCCAGCAGCTCATGTGAA	3401
Db	3597	GT	ACCACTTGT	CAAGCCGAACCTTAGCTGGGGTAAATACACAGGAGATGGTATGTGAA	3656
Qy	3402	CAG	CGCTGGCCCT	GAGTCAATCCGAACCTTTGACAGACAGTGAGACAAATGGAAGCACT	3461
Db	3657	CAG	GTATCCCCCT	GACTGCTGGTGGAAACCCCTACAGACAGTGAGACTGATGGAAGCACC	3716
Qy	3462	CC	TTAGCAGATT	TACACACTGCAGATGGGACACAGCAAGATGAACCAATTCAGACGCCAG	3521
Db	3717	CCG	TAGCCCACTTT	GACGCAACCCCAAGAAACGAGATTTGTGGAAATCCAT	3776
Qy	3522	GAC	AGTAAAGCC	ACTGCAGCTGTCCAGGAGTCAAGGTCACAGAAAGAGGCGGCTACT	3581
Db	3777	GAG	GAGAAATG	AGTGTGATCCCAAGTCAAGGGGCAAGAGACGAGGCGAGTCTCT	3836
Qy	3582	GCT	CAGAAAGAGG	AGCTTTCGACCTACCTAAATGTTCCAGGCCACAGGAAGAACATGGG	3641
Db	3837	GCA	CAGAAAGAG	AGGCGCTCCAGCACCTTCCAGTGTGTTGTTCCAGGAAGAACTAAAGAA	3896
Qy	3642	GA	GAAC	CAGGAAGAGATGTTCTTGAACTCAAGCAAGAGCTTACTGCTGAGCCGTG	3701
Db	3897	CA	ATTCAAAGAT	TGAAGACACTCTAGAGCATACAGATAAAGAGGTGTGAGTGGAAACTGTA	3956
Qy	3702	CCG	GTCTG	CAAGAGCTGAGGTGGTCAAGAGGGTGA- -----GGTTGACTGGTGTG	3752
Db	3957	TCC	ATTCTGT	CAAGAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCTGATGAGAAACC	4016
Qy	3753	GAT	GGAAAAAAG	GTCAAGAGAAACAGGAGGTGTTGTTGATCACTCTGG- ----ACCCAAC	3806


```

Oy 3300 ACTGACAGGCTTCTCCAGAGAGCTTG---AAGTTCTGAAAGTCAAGGACATGTA--- 3353
Db 3367 GTGGGGGAGACACCCCGAGAAAGCTTTGAAAAAGCTCTCAAGTCAAGAGACATAGAG 3426
Oy 3354 -----GACCATGTGCGCAGCTGCCA-----GGTTATCAAGCTCCAG 3389
Db 3427 TCCAGTGAAGCTTTGTAACCATTTGTCAAGCCGAACCTTAGTGGGGTAAATCAAGAG 3486
Oy 3390 CAGCTGATGAAACAGGCGGTGGCCCTGATCATCCGAAACCTTGACACAGATGAGACA 3449
Db 3487 ATGTGATGAAACAGGCTATCCCTGCTGATGAGAAACCCCTACAGACAGTGAAGT 3546
Oy 3450 AATGGAAGCATCTCTTTCAGATTTCAAGACATGACATGAGAGACACAGATGAAC 3509
Db 3547 GATGGAACACCCCGTATGCGACCTTTGACGACAGGACAAACCAAGAAACAGATTT 3606
Oy 3510 ATTGACAGCAGGACAGTAAGCACTGACGCTGTCAAGGACGTCAAGGTCAAGAGAA 3569
Db 3607 GTGGAATTCATGAGAGATAGGTCCATCTGGTACCACTGAGGGGGCAGAGAGCA 3666
Oy 3570 GAGCGGCTACTGTCTCAGAAAGAGAGCCCTTGACACTAATTAATGTTCCAGCCAG 3629
Db 3667 GAGGAGTTTCTGCAAGAGAAAGAGGCTCCAGACCTTCCAGTTTGTTCAGAGAA 3726
Oy 3630 GAAGAATGAGGAAAGAACAGAGAAAGATGTTCTTGAACCTACACAGACAGTACT 3689
Db 3727 GAAACTAAAGAACAAATCAAGATGAGAGACACTTACAGATACATTAAGAGGTCTA 3786
Oy 3690 GCTGAGCGGTGCGGCTTGGCAAGACTGAGGTGGGTCAAGAGGGGTAG----- 3740
Db 3787 GTGGAATCTGATTCATTTTGTCAAGACTGAGGGGACTCAAGAGGCTACAGATAGCT 3846
Oy 3741 GTTGACTGTTGATGAGAGAAAAAGTCAAGAGAAACAGAGGTGTTTGTACACTTGG- 3799
Db 3847 GATGAGAAAAACAAAGAGTACATTTTTCGAAGACTTGAAGGGGTCTATAGACACAGGC 3906
Oy 3800 -----ACCCAACATCTCAAAAGGCTGTGATGTGACATATGACATGATGAGAGTG 3854
Db 3907 ATACAGATCAGTGGGAAAAAGGTCACTGATGATGCTTAAAGGTGAAGGACAGAGAA 3966
Oy 3855 GCGGGGTGCAGAGAAAGAGAGTACTGAAGTGCAGAGT-----CTTACG 3899
Db 3967 GCTGAATGTAAGAGATATGCTCTTGAATGACAGAGTCAAGCTTAAGTCTCTCATCC 4026
Oy 3900 CTGAGAGAGAGAGATGAGAACTGACGTTGAAAAAGAGAAAAAGGAGACAAAGCCAGAG 3959
Db 4027 CCGGTGAGAGAGAGATGATGTTCAAGTCAAGGAGAGAAAAAGAGACAGAGCCAAACC 4086
Oy 3960 CAAGTGAAGTGAAG---AAGGTGACAGAGAAAGCCGCTCTCTGAGCATGAGAACTTAC 4016
Db 4087 CATGTGAATGAGAGAGACCTTGAACGAGAAACAGCTGTATCCGTATCTTAAGAGGTCACT 4146
Oy 4017 GGGAAAGCCAGTCTTCACTTGAACATGCCCAGCTCAGAGAGGGGAGAGGACCTGGGAAC 4076
Db 4147 AAGCAGCTCTCCAGACATGATGATGCCCCATATGATGAGGGGCAAGAGATCAGCAAT 4206
Oy 4077 CTGGAAGAGAGCCCTTCTC---TCCAGACCAAGCAAGAGAGGTGCAATAGAGTTCAA 4133
Db 4207 TTGGAAGAGAGCCCTCTCTCCCTGCTTAAGTCAAGAGAGAGAGATGACCAAAATTCAA 4266
Oy 4134 GTTCAAGAGCTTGACAAACAGTCACTCAAAACAGAGAGAGTGTGAAAAAGTCAAT---- 4189
Db 4267 GTTCAAGAGCTTGAGAGCATATCTTCAACAGCGCGCTGACAGAGAGAGAAAAAGGCTTAA 4326
Oy 4190 --AGAAACGGTTGTGATTTCAAGACAGGTGAAGTCAAGAGTGTAGGTGACACTTA 4247
Db 4327 GGAAGAACTGCAACATTTTGAAGACAGGTGAAGCTTGAAGCTCAGAGTGCATTTTA 4386
Oy 4248 TTACCAAGCTGAGAGATCTCTGCAACGGGTGGCCACTGGACTTTTCAGATGACAGAGAC 4307
Db 4387 GTTCTGGAAGAGAAATCTCTGAAAAAAATGAAGACTTTTGGCGCTATCCAGGGGAAAGAT 4446

```

```

Oy 4308 ACSTGACCCCTGGGCGCTGAGTCTCAGGACAGAAATCCATCCAAATCATATGAACTCTGCT 4367
Db 4447 GCTTGCCCAACAGGCGCCGACTGTCAAGCAAAATTCAGACACAGATAGATCTGCTACT 4506
Oy 4368 CTTGAAGACACCTTACATCTTACCTTACAGAGAAATTAAGCGATCCAGAGAGACGA 4427
Db 4507 ACCAAGAAAGGCTTAAATGTTCCGACTGGAAGAGAGAAAAACCATCATCATGAAGTGAAG 4566
Oy 4428 TCAGAGAGAGAGCAAGCCAGATGCTGTCTGATCTCATATCTTACAGGCAAGAGATACGA 4487
Db 4567 TCAGATGAAGTCAAGTGAAGAGGTTGCTTGGC-----AGAGGTCAAGAGTGTGCA 4620
Oy 4488 ATCAAAAAAGTCTCAAGGCTGAAGCTGAAGTCTGGAACCTTGAGAGTAAGAGCAAG 4547
Db 4621 ATTGAGAGAGATTTAGAGCTTGAAAAATGGATTTTGAACCTTGAGACCAAAAGCTAATA 4680
Oy 4548 ATTGTGCTGAAGTCAATTCAGACAGCCGTTGACAGTTGCAAGCTAC--AGAAACAGCC 4604
Db 4681 CTTGTCCAAAACATCATTCAGACAGCCGTTGACAGTTTGTACGTACAGAAAGAAACAGCC 4740
Oy 4605 CCCGAATCTAGCTTATGATTCACAGACCCAGGTTCTGATCAGAGCTTGAACAGAG 4664
Db 4741 ACCGAAATGTTGAGTCTGAGTTACAGACCAAGCTCAGGTATAAAGCTGACAGCCAG 4800
Oy 4665 GA 4666
Db 4801 GA 4802

```

```

RESULT 6
US-09-724-676A-12736
; Sequence 12736, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent version 3.2
; SEQ ID NO 12736
; LENGTH: 5773
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (91)..(91)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12736

```

```

Query Match 31.2%; Score 1920.8; DB 5; Length 5773;
Best Local Similarity 68.2%; Pred. No. 0;
Matches 3072; Conservative 0; Mismatches 1247; Indels 183; Gaps 22;

```

```

Oy 227 GTTGACAGCGAGAGTCAAGATGTGAGAGAAAAAGACGAGTTGAAGAAATGGCGGCC 356
Db 332 GTTGACAGAGAGACTCTGAGAGATGTGAGCAAGAGAGCTCCGATTAAGAGATGGCTACT 411
Oy 357 AACTCAAGCTGTGAGATATCAAAAGAGATGGGACAGAGAGACATCAAGAAATAATT 416
Db 412 AAGTCAAGGGTGTTCACAGACATCAAGATATGGGACAGAGAGACACCCGAATAATTC 471
Oy 417 GAACGATCCCTGCTTCAGAAAACAATGTGAAGAAATGTACAGCTCTGATGCCAG 476
Db 472 GAACGATTCCTTCTTCAGAAAGCAATTTAGAAAGCTTACACACCACTGAGTCCAG 531
Oy 477 GCTAATGATGTGGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTCACGGTGAAG 536
Db 532 GCTAATGATATTTGATTTAAGAGAGGTGTTTAAGTTTGTGGCTTTAATTCATGTGAAA 591
Oy 537 AAGATTAAGAAATGAAGATCATCTGTCCAATCTCACTGTCCAAGAGATGAAGGC 596
Db 592 AAGATTAAGACAGAGAGCTGACATGTCCAGCTACCTGTGAAAGAAAGATGAAGGG 651

```

QY 597 GAAGGGCAGAGAGCTCTGTCCGAGCTGGAGACCAACAGAGAGCCAGTGTGGAGACTGCC 656
Db 652 GAGGAGCAGCA-----GGGGCTGGCACCACACAGGACCCACAGCTTGGGGCT--- 699
QY 657 GTCCGAGAGTCAGCATCCAAAGAAAGTGTAGCTGAAGCAATCCACAGAGAGCAAGAGGC 716
Db 700 ---GGAGAGCAGCATCAAAAGAAAGCAACCCAAACAATCTACAGAGAAACCCGAGAG 756
QY 717 ACCCTGAAGCAAGAACAGAGCAGACAGAAATCCCCCTTCAAGCCGAATCTGTATCAAGCG 776
Db 757 ACCCTGAAGCGTGAGCAAGCCACGCAAGAAATTTCTCCCCACGCGCAATCTGGCCAA--- 813
QY 777 GCTGAGAGAGCCAAAGATGAAGAGAGAGAAAGAAACAGAGAAAGAGCCACCAAGTCC 836
Db 814 GCAGTGGAGGAATCAAAAGAGGAAGGAGAGAGAAACAGAGAAAGAAACCTTAGCAAGTCT 873
QY 837 CCAGAAATCCCGAGCAGCCAGTCAACAGTGAGACACATCTTCTTCAAGAGTCTTCTC 896
Db 874 GCAGAAATCTCCGACTAGTCCCGTGACCAAGTGAAACAGGATCAACCTTCAAAAAATTTCTC 933
QY 897 ACTCACGTTGGGCGGCTGGCGCAAGAGACCAAGCTTCAAGAAATCAAAAGAGGATCAT 956
Db 934 ACTCAAGTTGGGCGGCTGGCGCAAAAGACCAAGTTCAGAGAGCCGAGGAGATGAA 993
QY 957 CTGGAATCTGCCGAGAGAGAAAGGAGCAAGAGCAGAAAGTAGAGAGAGAAAG 1016
Db 994 GTGGAAGCTTCAGAGAGAGAAAGGAGCAAGAGCCAGAGAAAGTAGACACAGAGAAAGAC 1053
QY 1017 GAAAGACAGAGCAGCTCGAGGA-----CGAGAG 1049
Db 1054 GAAAGGAGAGGTTGGCTTCGAGAAACTGACCGCTCCGAGCAAGCCACCCACAGGAG 1113
QY 1050 CCGCAGAGAGACACAGACCAGGCCAGTGTGTACAGACACTACAGAAAGGTGGAGCTGCCT 1109
Db 1114 CCGCAGAAAGTGCCACGAGCCCGGTTATCAGCTGATATGAGAAAGTTGAGCTGCC 1173
QY 1110 TTGGAAGACCAAGTTGTGACCTGGAGCATCGTACAGAGGAAGTGTGCTCTTTGGCA 1169
Db 1174 TCAGAGGAGCAAGTCAGTGGCTCGCAGGACCTTCTGAAGAGAAACCTGCTCCGTGGCG 1233
QY 1170 ACCGAAGTGTGTATGAGAGATGGAGGCCACCAAGAA---GTTGTTGAGAGGTCCAC 1226
Db 1234 ACAAGTGTGTATGAGAAATAGAAAGTCCACCAAGAGAGGTTGTGGCCGAAAGTCCAC 1293
QY 1227 GTGAGCACCGTGGAGAGACAGAGAGAGCAGGAGGAGGAGGAGGCTGAAGGGGC 1286
Db 1294 GTACACCGTGGAGGAGAGAACCGAAGACAGAAA----- 1329
QY 1287 GTGTGTGTAAGGAACAGGAGAAATCTTGGCCCTCAGAGAACTGGGTGAGCCCCAGGAG 1346
Db 1330 ACGAGGTGGAAGAAACAGCAGGGTCTGTGCCAGCTGAAGAAATGGTTGAAATGGATGCA 1389
QY 1347 GTCCCCAGGAGCTGAGCCTGTGAGGAGCTGATGAGAGCAGAGAGATGTGTCTCT 1406
Db 1390 GAACCTCAGGAAGCCGAACCTGCCAAGAGCTGTGAAGCTCAAGAAACGTTGTGTTTC 1449
QY 1407 GGAGGAGACCACTCAACTGACAGACTAGTCTTGAAGAGAGAGCGTSCCCAAACAC 1466
Db 1450 GGAGAGNCCCTACACAGGAGCTGACCTCAGTCTCTATGAGAAAGTGTCTCCAAACCC 1509
QY 1467 CCAGAAAGCAATGTTCAGTGAAGGTGGAGATGCTGCTCTCAGGAAAGAAATCAAGGTACAG 1526
Db 1510 CCCGAAGCGGTTGTGAGTGAAGTGGAAATGCTGTCATCACAGGAGAGAAATGAAGTGCAG 1569
QY 1527 GGAAGTCCCTTGAAGAACTCTCAGTAGCTCAGGCTTAAAGAGCTGTCTGGAGAAAG 1586
Db 1570 GGAAGTCCACTAAAGAGCTTTTACCAGCACTGGCTTAAAAAGCTTTCTGGAAGAAA 1629
QY 1587 CAGAAGGGGAAACAGGAGGTGGGGAGACGAAGAGCTTGAGAAATACCAACATTCAC 1646
Db 1630 CAGAAGGGAAAGAGGA---GGAGGAGACAGGAATCAGGGGAGCACTCAGGTTCCTCA 1686

QY 1647 ACCGAATCCCAAGAGAGTGTGTATGAGCAGAGAGGAGAGAGCTCTGTGCTGTCCCGAG 1706
Db 1687 GCCGATTTCTCCGACACCCAGGAGAGCAAAAGGGGAGAGCTCTGTGCTCATCCCTGAG 1746
QY 1707 GAGCCTGAGGAGACCAAGTGTCTGGAGAAAGGCGCTGGAAGCAACCCAGGATGGGAA 1766
Db 1747 GAGCCGAGAGATCAGTGTCTGGAAAGGGCTTAGCCGAGGTGCAGAGGATGGGAA 1806
QY 1767 GCTGAGGAAGAACTACTTCCGATGGAGAGAAAGAGAGAGGAGTCACTCCCTGGGCA 1826
Db 1807 GCTGGAAGAAGGAGCTACTTCCGATGGAGAGAAAAAAGAGAAGTGTCACTCCCTGGCA 1866
QY 1827 TCCTTCAAAAAGATGTGACACCCCAAGAAACGGGTCCGAAGACCTTCTGAGAGTGAAG 1886
Db 1867 TCATTCAAAAAGATGTGACGCCCAAGAGCGTGTAGACGGCTTTCGGAAGTGAATAA 1926
QY 1887 GAGNAGAGCTGAGAGAGTCAAGAGCGCCACCTGTGCTCTCACTGATAGCACAGTGTCA 1946
Db 1927 GAAGATGAGCTGGAAGAGTCAAGAGCGCTACCTTGTCTTCCACCGAGAGCACAGCTCT 1986
QY 1947 GAAATCCAAGATGAAGTCAAAACTGTTGTGAGGAAACAAAGCCAGAGGAACCAAGCGT 2006
Db 1987 GAAATGCAAGAAAGATGAAGGGAGCGTGAAGAGCCAAAGCCGGAAGAACCAAGCGC 2046
QY 2007 AGGTGATACCTTCAAGTGTCTTGGGAAGCACTGATTTGTGTCCGATCATCCAAGAGAGA 2066
Db 2047 AAGGTGATACCTTCAAGTGTCTTGGGAAGCTTTAAATTTGTGTGGATCATCCAAGAAA 2106
QY 2067 GCAAGNAGGATCTCTTTCAGATGATGAGAGGGCCAGGACACTGGAGGGGACAGT 2126
Db 2107 GCAAGGAGAGGGTCTCTTCTGATGAGAAAGGGGAGCAAAAGCAATGGAGGAGACAC 2166
QY 2127 CACAGAGCAGAGAGGCGCAAGAAAGCCGGAACAGACAGCTGTTCCTGCCAGC 2186
Db 2167 CAGAAAGCTGATGAGCGCGAAAGACAAAGAGACGGGACAGACGGGATCTTGTCTGT 2226
QY 2187 ACCGAGAGCAGACCAAGCGCAAGGAAGTTCCTCAACCGAGCCAGCGGAAGCCCTTCC 2246
Db 2227 TCCCAAGAACATGATCAGGGCAGGGAAGTTCCTCCCGGAGCAAGCTGGAAGCCCTACC 2286
QY 2247 GAAGGGAGGCTCTCCACTTGGGAGTCAATTTAAAGATTTAGTCACTCCAAGAAAAA 2306
Db 2287 GAAGGGAGGCGGTTCCACTCGGAGTCAATTTAAAGGTTAGTCAGCCCAAGAAAAA 2346
QY 2307 TCCAAAGTCAAAAATGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAG 2354
Db 2347 TCAAAGTCCAAGCTGGAGAGAAAGCGAAGACTCCATAGCTGGTCTGTGTAGAACAT 2406
QY 2355 TTGTCCACTGAGATCGAACCGAGTAGAGAAATCTTTGGGTTTCCATTAAGAAATTCATC 2414
Db 2407 TCCACTCCAGACACTGAACCCGGTAAAGAGAAATCTGTGGTCTCAATCAAGAAAGTTATT 2466
QY 2415 CCGGAGCGGGAAGAAAGGGCAGACGGGAGCAAGCAAGCAAGCACTGTGGAGACTCA 2474
Db 2467 CTTGGAGCAAGGAGAAAGGGCCAGATGGGAAACAAAGCAAGCCCTGTGTAAGACGCA 2526
QY 2475 GGGCCAGTGGAGATAATGAGGACGACCTTAATGTCCAGCCGTCTGTGCTCTGTGAG 2534
Db 2527 GGGCCACAGGGGCCAACAGAGATGACTCTGATGTCCCGCCGTGTGCTCTGTCTGAG 2586
QY 2535 TATAATGCAAGTGGAGGGGAGAAAGATGG-----AAGCCAGGGGAAATACGGAGCTG 2585
Db 2587 TATGATGCTGTAGAAAGGGAGAAAATGGAGGACAGCAAGCCCAAAAGCGCAGAGCAG 2646
QY 2586 CCCAGGTGCTGGGGCTGTGACGTGTCGAGGAGTCTGTAAGACTCTGTGTCACACT 2645
Db 2647 CCCAGCAGAAAGCAGCCACTGAGGTGTCCAGGAGCTCAGCGAGAGTCAAGTTCATG 2706
QY 2646 GTGAGTGTCCAGTCAATTTGATGGACAGGGCAGTCAAGTGTGTAAGAGCGGTCTCCT 2705
Db 2707 ATGGCAGCAGCTGTGCTGACGGGACGAGGGCAGTCACTTATTTGAAGAAAGTCTCT 2766
QY 2706 TCGTGGATATCCGCTTCCGTAACAGAAACCTCTTTGAACACACAGCGGGAGAACCATGCCA 2765

```

Db 2767 TCTTGATATGTCTTCAGTACAGAACCTCTTGAAACAGTAAAGCTGAAAGCCCACTG 2826
Oy 2766 CCTGTGAAGAGTCACTGAAAAAGACATCATTTGCA---GAAGAACTCTGTGTCTCAC 2822
Db 2827 TTTACTGAGAGTATTGGAAGAGAAATTTGCAAGAAAGAACCCCAAGTTACT 2886
Oy 2823 CAGAGCTTACAGAGGCTTAAAGATGCCATGACATGTGTACCAAGTGAAGTGAATTC 2882
Db 2887 GAACCTCTGCAGAGAACAGAGAGGCCCGGCGCAGACAGTGTGTAGTAGAGCCGAAATTG 2946
Oy 2883 ACCCTGAGAGCTGTGACAGCCACAGACCTCAGAGGCTCTCCGACTGAAAGATTAC 2942
Db 2947 ACCCTCGAAGCTGTGACAGCTGTGCAAGAACTGCAAGGCCATTGGGTGCCGAAAGAAC 3006
Oy 2943 GAAGCATCGGGGCCGAAAGACACAGACATGTGTCTCCAGTTTCCAGCTGACTGAC 3002
Db 3007 GAAGCATCTGCTGTGAGAGAGACACAGAAATGTGTGACGAGCTCTCCAGTTAACGAC 3066
Oy 3003 TCCCCAGACACCAAGAGAACCCCAAGTTCAAGAGTGAAGAGTGTGTCTAGAT 3062
Db 3067 TCCCCAGACACCAAGAGAGGCCACTCGGTGACAGAGTGAAGTGTGCGTACCTGAC 3126
Oy 3063 ACAGAAAGAGAGAGGCCGACAGCGAGGCCATCTCCAAAGCGGTGACAGACAGGTGAAA 3122
Db 3127 ATGAAAGAGCAAGAGAGCGGACTCAAGAGTCTCTCAGGCAGTGCGAGAAAGGTGAAA 3186
Oy 3123 GAGGAGTCCAGAGTCTCTGCAACCA---GAGTGTGACAGAAACGGGCTCAAAACACTG 3179
Db 3187 GAGGAATCCCACTGCTGTGACCGGTGGCCAGAAAGATGTGCTTCAAGCTGTGTGACAGA 3246
Oy 3180 GAGAGCTTGAAGAGTGAAGAGAGACCTCCGAAGTGTGCTGTGAGAGAAAGAGAC 3239
Db 3247 GCAGAGGCGAAGAAAGACAGAAAGAGCGTGAAGCGTCTCAAGAAAGAGACGAGAT 3306
Oy 3240 GTTATGCCGAAGAGACCGGTGAGAGAGTGAAGTGAAGTCTTGTGACAGGCTCTGAG 3299
Db 3307 GTAGGTGTGAAGTGAATGCTCAGAGAGCAAAACCTGTAACCTTTTACAAAGGAAAGGTG 3366
Oy 3300 ACTGACAGGCTACTCTCAGAGAGCCTTG---AAGTTCTGAAAGTCAAGCAGATGTA--- 3353
Db 3367 GTGGGGCAGACCAACCCAGAAAGCTTTGAAAAAGTCTCTCAAGTCAAGAGACATAGAG 3426
Oy 3354 -----GACCATGTGCGCACGTGCCA-----GGTTATCAAGTCCAG 3389
Db 3427 TCCAGTAGCTGTGAACCACTGTCAAGCCGAAACCTTAGCTGGGGTAAATCAACAGAG 3486
Oy 3390 CAGCTGATGGAACAGGCGTGGCCCTGTGATCCGAAACCTTGAACAGAGTGAACA 3449
Db 3487 ATGTGTATGGAACAGGCTATCCCTCTGACTGTGTGAAACCTTACAGAGTGAAGT 3546
Oy 3450 AATGGAAGCACTCCCTTAGACATTCAGACACTGACAGTGGGACACAGCAAGATGAAAC 3509
Db 3547 GATGGAAGACACCCCGTAGCCGACTTTGACGACACAGGACAAACAGAAAGACAGAT 3606
Oy 3510 ATTGACAGCAGAGACATGAAGCCCACTGACGTGTCAAGGACGTCAAGGTCAACAGAAA 3569
Db 3607 GTGGAATTCATGAGAGATGAGTGCATCTGTGATCCAGTCAAGGCGGACACAGAAACA 3666
Oy 3570 GAGGCGGCTACTGCTCAGAAAGAGAGCCTTGCACATCACTAAATTAATTTCCAGCCAG 3629
Db 3667 GAGGCACTTCTGTGACAGAAAGAGGCGCTTCAAGCACTTCCAGTTTGTGTCCAGAAA 3726
Oy 3630 GAAGAACATGGGGAGAGAACAGAGAGATGTTCTTGAACCTACACAGCAAGAGTACT 3689
Db 3727 GAAACTAAAGAACATCAAGATGAAGACCTTAGAGCATACAGATAAGAGGTGTCA 3786
Oy 3690 GCTGACAGCCGTCCTGTTGCAAAAGTGAAGTGTCAAGAGGTGAG----- 3740
Db 3787 GTGGAACACTGTATTCATCTGTCAAAAGTCAAGAGGAGTCAAGAGGCTGACAGATATGCT 3846
Oy 3741 GTTGACTGTTGATGAGAAAAAGTCAAAAGAAACAGAGGTGTTTGTACACTCTGG- 3799

```

```

Db 3847 GATGAGAAAACAAAGACGTACCACTTTTTCGAAAGACTTGAAGGGGTCTATAGACAGAGC 3906
Oy 3800 -----ACCCAAAGTCAAAAGGCTGCTGATGTGACATATATGACGTGAAGTGGAGTG 3854
Db 3907 ATTAAGTCAAGTCGGGAAAAAGTCACTGAAGTGTGCTTTAAAGTGAAGGACAGAAAGAA 3966
Oy 3855 GCCGGGTGTCAAGGAAAAAGAGTACTGAAGTGCAGAGT-----CTTAGC 3899
Db 3967 GCTGAATGTAAAAAGATGATGTCTTTGAACCTGACAGGTACAGCTAAGTCTCTTCAATCC 4026
Oy 3900 CTGAGAGAGGAGATGAGAACTGACGTTGAAAAGAGAGAAAAAGGAGAACAAAGCCAGAG 3959
Db 4027 CCGTGGAGAGAGATGTGATGTTCAAGTCGAAAGGAGAAAAAGAGAGAGAGCCAAAC 4086
Oy 3960 CAATGTAGTGAAG---AAGTGAACAGAAAAAGCCGCTCTGTAGCATGAAAGAACTTAC 4016
Db 4087 CATGTGAATGAAGAAAGCTTGAACAGAAACAGCTGTATCCGTATGTGAAGAGTGAAGT 4146
Oy 4017 GGAAGCAAGTCTGACACTTGAACATGACCCAGCTCAGAGAGGGGGAAGGCACTGGAAAGC 4076
Db 4147 AAGCAAGTCTCTCAGACAGTGAATGTGCCATATCATATGAGGCGCAAGAAAGTACAGAT 4206
Oy 4077 CTTGAGAGAAAGCCCTTCTC---TCCAGACCAAGACAAAGCAGGTTGCATAGAGTTCAA 4133
Db 4207 TTGGAAGAGAACCTCTCTCCCTGCTTAGGTCAAGAGAGGAGATGACCAAAATTTCAA 4266
Oy 4134 GTTCAAAAGCTGACACACACATCTCAAAACAGACAGAGCTGTGAAAAGTCAAT--- 4189
Db 4267 GTTCAAGAGCTGTGAGCATCATTCATCTTAAACAGCGGCTGACAGAGAGAAAGTCTTA 4326
Oy 4190 --AGAAACGTTGTGATTTCAAGACAGGTGAAGTCCAGAGTGTGTAGGTGACACATTA 4247
Db 4327 GGAAGAACTGCAACATTTTGAAGAAAGGTGAACGTTGAGCCTGCAAGGTGCACATTTA 4386
Oy 4248 TTACACAGTGAAGAGTCTCTGCAACCGGTGCGCACTTGACATTTCAAGTGCAGAGAC 4307
Db 4387 GTTCTGAGAGAGAAATCTCTGAAAAAATGAAGACTTTGCCGTATCCAGGGAAAGAT 4446
Oy 4308 ACGGTACCCCTGGGGCTGAGTCTCAGGACAGAAATCCATCCCATCATATGTAATCTCTGCT 4367
Db 4447 GCTGTGCCCAAGGGCCCGACTGTCAAGCAAAATCGACACAGATGATGTAATCTGACT 4506
Oy 4368 CCGTGAAGCAACCTTACATCCGACCTTCAAGGAGAAATTAAGCGCATCCAGAGAGAGCA 4427
Db 4507 ACAGAAAGAGCTTTAAGTTCGACCTTGAGAGAGAGAAACCATATCACTGAATGAGAG 4566
Db 4567 TCAATGTAAGTGAATGAGACAGATGTTGCTTGGC-----AGAGGTCAAAAGTGAAGTGA 4620
Oy 4488 ATCGAAAAAGTCTCAGAGGCTGAACCTGAGATCTCGAACTTGAAGATGAAGCAACAG 4547
Db 4621 ATTGAAGAGATTTAAGCTGAAATGGAATTTTGAACCTTGAAGCCAAAGAGATGAAA 4680
Oy 4548 ATTTGTGCAAGTCATTCAGACAGCGGTTGACAGATTCGACAGTAC---AGAAACAGCC 4604
Db 4681 CTGTGCAAAACATCATCAACAGCGGTTGACAGATTTGATGACAGAAAGAGCC 4740
Oy 4605 CCGGAACTCATGTCTTATGATTTACAGACCAGAGTTCCTGCATCAGGCTTGAACAGCAG 4664
Db 4741 ACCGAAATGTTGACGTCTGAATTTACAGACACAAAGCTCAAGTAAAGCTGACAGCAG 4800
Oy 4665 GA 4666
Db 4801 GA 4802

```

```

RESULT 7
US-10-152-319A-1312/c
; Sequence 1312, Application US/10152319A
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark

```



```
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1312
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. A1233818
; US-10-152-319A-1312

Query Match      8.9%; Score 550; DB 6; Length 563;
Best Local Similarity 99.1%; Pred. No. 3.1e-120;
Matches 553; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5601 ATACGCGAGCTTTGAAAGCTCCAAAGCTCAAGCTCAACTGTGTAACCTGCGAGCAACAGATAACATT 5660
DB 563 AAAGCGAGCTTTGAAAGCTCCAAAGCTTCAACTGTGTAACCTGCGAGCAACAGATAACATT 504
QY 5661 CCTGGCAAGAGACAGACAGCTCTTTTAAAGTTTACTGATGCTTAGATCTGTGGGCTTCT 5720
DB 503 CCTGGCAAGAGAAACAGACAGCTCTTTTAAAGTTTACTGATGCTTAGATCTGTGGGCTTCT 444
QY 5721 AGTCCTCTGAAAGTGGTGTGTTTCTTATGACAGCGAGCTCAGAAATAAAGCAACCCCATTT 5780
DB 443 AGTCCTCTGAAAGTGGTGTGTTTCTTATGACAGCGAGCTCAGAAATAAAGCAACCCCATTT 384
QY 5781 TGAACAATCCAGGATGCCCCAATAATACCATGATTTTTTCCCCCTTTTGTCTTAATCCAG 5840
DB 383 TGAACAATCCAGGATGCCCCAATAATACCATGATTTTTTCCCCCTTTTGTCTTAATCCAG 324
QY 5841 TCCAGGTTGGAAGAAGTCTCTCTGTGTCAGATTAAGCCCTGTCTCTTAATGATATGGA 5900
DB 323 TCCAGGTTGGAAGAAGTCTCTCTGTGTCAGATTAAGCCCTGTCTCTTAATGATATGGA 264
QY 5901 CAATGATGTGCTCAAGGCCATGAGATGTTTCTTAATGCAAGGAATCTGTTGTACGT 5960
DB 263 CAATGATGTGCTCAAGGCCATGAGATGTTTCTTAATGCAAGGAATCTGTTGTACGT 204
QY 5961 TTTTGTGATGTACTCTTCTAATGCGACCGAAATTCATATGACATCGAAGTGAGTCTGT 6020
DB 203 TTTTGTGATGTACTCTTCTAATGCGACCGAAATTCATATGACATCGAAGTGAGTCTGT 144
QY 6021 TTCTTTACAGATGATTTTATGATAGACTGGAGTTGTCTGTGTTATATCTGTGCCCT 6080
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1312
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. A1233818
; US-10-152-319A-1312

Query Match      2.6%; Score 162.2; DB 6; Length 178;
Best Local Similarity 97.8%; Pred. No. 8.3e-29;
Matches 175; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 5962 TTTTGTGATGTACTCTTCTTATGCTGGACCGAAATTCATATGACATCGAAGTCACTCCTGT 6021
DB 1 TTTTGTGATGTACTCTTCTTATGCTGGACCGAAATTCATATGACATCGAAGTCACTCCTGT 60
QY 6022 TCCTTACAGATGATTTTTCATAGATACGAGTTTGTCTGTGTTATATCTGTGCCCTT 6081
DB 61 TCCTTACAGATGATTTTTCATAGATACGAGTTTGTCTGTGTTATATCTGT-CCCTTT 119
QY 6082 CTTTAAAGCAAAATGTTGCATTATGTTCTCTTGGATAAAATGTTGATTTTGACAACTGATTT 6140
DB 120 CTTTAAAGCAAAATGTTGCATTATGTTCTCTTGGATAAAATGTTGATTTTGACAACTGATTT 178
```

```

RESULT 9
US-10-152-319A-1193/c
; Sequence 1193, Application US/10152319A
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Potter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; PRIOR FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1193
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AI228494
US-10-152-319A-1193

Query Match
Best Local Similarity 83.0%; Score 124.6; DB 6; Length 180;
Matches 142; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 5987 GACCGAATTCATATGCAATGCAAGTACGCTGTTCTTTTACAGATGTAATTTGATAGA 6046
DB 180 GGCCGAATTAATTAATGAGATGAGAGTCTGTTTCTTTTACAAAAGGAGATTTGAAAAA 121
QY 6047 TACTGAGATTTGTCTGTATATCTGTGTCCTTCTTTAACAACAATGTCATTATGT 6106
DB 120 AACTGGGGTTGGTGGTAAATTTGGGCCCCCTTTTAAAAACAATTTGCTTAAGT 61
QY 6107 TCCTTGATAATGATGATTTGACACATGATTTAAATTAACAATTTGAC 6157
DB 60 TCCCTTGAAAAATGGAGTTGGCACTGATTTTAATTAACAATTTGCC 10

RESULT 10
US-09-724-676-12734
; Sequence 12734, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```

```

; SEQ ID NO 12734
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12734

Query Match
Best Local Similarity 64.7%; Score 119; DB 5; Length 753;
Matches 214; Conservative 0; Mismatches 105; Indels 12; Gaps 2;

QY 3 GGCCGGCGGAGTAGAAGCACTGAGCCATGGCGCAGAGATTTCACAGACAGCG 62
DB 162 GGCTAGGCGGGAGAGAGTCCGAGAGCCATGGGCGCGGAGCTTCACAGACAGCG 221
QY 63 AGCCCGACGACCGCG---GGGGGAGCGACACCGCCAGCGAGCTGTCTCAGTGGCCAT 119
DB 222 AGCCCGAGAGAGCGCGCGCGGAGGAGCTTCACCGCTAGCGCCGAGCGGCGGCG 281
QY 120 GGGCCCGAGCTGAAGCTCGGGA-----GCAGCTGAGACCCCGCGACCGGAGC 170
DB 282 GGCCCTCGGCGGAGGCGGCGCGACACACCGGACCCCGCATGCTGCTCGGAC 341
QY 171 CCCGCAACCAAGCTCCACAGAAAGATGCGCAGCTGTTCTGTCAACGGCGTAGTGA 230
DB 342 CCCGCAACCAAGCTCCTACAGAAAGATGCTGCTGCTCAACATCAATGGGTAGTGA 401
QY 231 CAAGAGATGTCATGTCACAGAGAAAACGAGAGGGGCGAGAGAAAGTGTGAT 290
DB 402 CAAGATGAGCTCAAGCTTCAGAGAGGTGACTTAATAGGCGAGAAAGAGCCCTGA 461
QY 291 GAGGATGTTGACAGCGAGAGTCAAGATG 321
DB 462 CAAGAGAGCTTAACAGCGAGAGAAAG 492

RESULT 11
US-09-724-676A-12734
; Sequence 12734, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12734
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12734

Query Match
Best Local Similarity 64.7%; Score 119; DB 5; Length 753;
Matches 214; Conservative 0; Mismatches 105; Indels 12; Gaps 2;

QY 3 GGCCGGCGGAGTAGAAGCACTGAGCCATGGCGCAGAGATTTCACAGACAGCG 62
DB 162 GGCTAGGCGGGAGAGAGTCCGAGAGCCATGGGCGCGGAGCTTCACAGAGAGCG 221
QY 63 AGCCCGACGACCGCG---GGGGGAGCGACACCGCCAGCGAGCTGTCTCAGTGGCCAT 119
DB 222 AGCCCGAGAGAGCGCGCGCGGAGGAGCTTCACCGCTAGCGCCGAGCGGCGGCG 281
QY 120 GGGCCCGAGCTGAAGCTCGGGA-----GCAGCTGAGACCCCGCGACCGGAGC 170
DB 282 GGCCCTCGGCGGAGGCGGCGCGACACACCGGACCCCGCATGCTGCTCGGAC 341
QY 171 CCCGCAACCAAGCTCCACAGAAAGATGCGCAGCTGTTCTGTCAACGGCGTAGTGA 230
DB 342 CCCGCAACCAAGCTCCTACAGAAAGATGCTGCTGCTCAACATCAATGGGTAGTGA 401
QY 231 CAAGAGATGTCATGTCACAGAGAAAACGAGAGGGGCGAGAGAAAGTGTGAT 290

```


Db	7357	GGCAGGACGAGGAGGGCGAGGACGAGGGCGAGGACGAGGAGTGGAGGCCGGG	7298
Qy	1606	GTGGGGGAGCAGAGAGCCTGGAGAAATCCACATTTACACCGAATCCCCAGAGAGTG	1665
Db	7297	TCGAGGAGGCAGTGGAGGCCGGGTCGAGGAGGTAGTCGAGGCCGGGTCCGAGGAGGTAG	7238
Qy	1666	CTGATGAGCAGAGGGAGAG	1685
Db	7237	TGGAGGCCCGGGGTAGAG	7218

Search completed: December 13, 2002, 00:06:21
Job time : 448 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:12:21 ; Search time 329 Seconds
(without alignments)
3127.641 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGNSTGTEQSPQAGSDT.....AWAQRKCLPRLQLKAPVSK 1596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main.*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8073	100.0	1596	1 PCT-US97-06830-4	Sequence 4, Appli
2	8073	100.0	1596	13 US-08-665-401-4	Sequence 4, Appli
3	8073	100.0	1596	13 US-08-978-277A-4	Sequence 4, Appli
4	8073	100.0	1596	23 US-09-902-432-4	Sequence 4, Appli
5	5421.5	67.2	1346	1 PCT-US97-06830-2	Sequence 2, Appli
6	5421.5	67.2	1346	10 US-08-665-401-2	Sequence 2, Appli

7	5421.5	67.2	1346	13	US-08-978-277A-2	Sequence 2, Appli
8	5421.5	67.2	1346	23	US-09-902-432-2	Sequence 2, Appli
9	4318.5	53.5	1782	27	US-60-389-987-2097	Sequence 2097, Ap
10	4318.5	53.5	1782	21	US-60-412-418-2097	Sequence 2097, Ap
11	4294	53.2	1781	21	US-09-738-877-3	Sequence 3, Appli
12	4294	53.2	1781	23	US-09-961-403-13	Sequence 13, Appli
13	4262.5	52.8	1795	1	PCT-US01-08631-51377	Sequence 51377, A
14	4253	52.7	1779	27	US-60-245-228-307	Sequence 307, App
15	780	9.7	732	21	US-09-758-442-562	Sequence 562, App
16	780	9.7	732	26	US-10-217-623-562	Sequence 562, App
17	752	9.3	212	27	US-60-182-569-13182	Sequence 1318, Ap
18	642	8.0	2768	20	US-09-614-150-31983	Sequence 31983, A
19	642	8.0	2768	27	US-60-173-464-24710	Sequence 24710, A
20	642	8.0	2768	27	US-60-191-637-31555	Sequence 31555, A
21	610.5	7.6	174	27	US-60-182-093-2249	Sequence 2249, Ap
22	561	6.9	194	27	US-60-186-662-1015	Sequence 1015, Ap
23	557.5	6.9	6358	27	US-60-191-637-27445	Sequence 27445, A
24	557.5	6.9	6358	27	US-60-191-681-22054	Sequence 22054, A
25	557.5	6.9	6815	20	US-09-614-150-27225	Sequence 27225, A
26	541.5	6.7	1770	27	US-60-196-710-5536	Sequence 5536, Ap
27	536	6.6	144	27	US-60-178-307-2458	Sequence 2458, Ap
28	534.5	6.6	169	27	US-60-196-710-6834	Sequence 6834, Ap
29	493	6.1	3924	19	US-09-538-092-1246	Sequence 1246, Ap
30	492	6.1	2468	19	US-09-538-092-1135	Sequence 1135, Ap
31	492	6.1	2468	23	US-09-976-594-726	Sequence 726, App
32	492	6.1	2519	1	PCT-US01-08631-46995	Sequence 46995, A
33	489	6.1	6642	27	US-60-360-039-5013	Sequence 5013, A
34	489	6.0	3263	20	US-09-614-150-28422	Sequence 28422, A
35	485	6.0	3267	27	US-60-191-637-27442	Sequence 27442, A
36	485	6.0	3267	27	US-60-191-637-27442	Sequence 27442, A
37	485	6.0	3267	27	US-60-191-637-27442	Sequence 27442, A
38	485	6.0	3267	27	US-60-191-637-27442	Sequence 27442, A
39	484	6.0	1778	27	US-60-188-162-4950	Sequence 4950, Ap
40	473.5	5.9	8991	11	US-08-714-741-32	Sequence 32, Appli
41	470	5.8	2364	26	US-10-205-331-66	Sequence 66, Appli
42	465.5	5.8	212	27	US-60-196-710-4143	Sequence 4143, Ap
43	464.5	5.8	2024	27	US-60-167-217-5735	Sequence 5735, Ap
44	464.5	5.8	2024	27	US-60-173-464-4625	Sequence 4625, Ap
45	462	5.7	1786	13	US-08-973-462-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
PCT-US97-06830-4
; Sequence 4, Application PC/TUS9706830
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06830
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/665,401
; FILING DATE: 18-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154


```

REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1596 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
PCT-US97-06830-4

Query Match      100.0%; Score 8073; DB 1; Length 1596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAGSTORSEPOAGSDTPSELVLSGHGPAEASGAAGDPADADPATKLPQKNGQLSS 60
Db 1 MGAGSTORSEPOAGSDTPSELVLSGHGPAEASGAAGDPADADPATKLPQKNGQLSS 60
QY 61 VNGVAGQGDVHVOEENOGQEEVVDVQRESESDVREKDRVEEMANSTAVEDITKDG 120
Db 61 VNGVAGQGDVHVOEENOGQEEVVDVQRESESDVREKDRVEEMANSTAVEDITKDG 120
QY 121 QETSEIIEQIIPASBNNEEMVQPAESQANDVGFKKVFQFKTKVKDKNEKSDTVQL 180
Db 121 QETSEIIEQIIPASBNNEEMVQPAESQANDVGFKKVFQFKTKVKDKNEKSDTVQL 180
QY 181 LTVKDEGEGAGASVAGADHOPSVETAVGESASELSKOSTEKOGTLKQEGSTEIP 240
Db 181 LTVKDEGEGAGASVAGADHOPSVETAVGESASELSKOSTEKOGTLKQEGSTEIP 240
QY 241 LOAESDOAAEEBAKDEGEKQKEPTKSPSPSPVNSSETTSSFKKFTTHGAMRKKTSS 300
Db 241 LOAESDOAAEEBAKDEGEKQKEPTKSPSPSPVNSSETTSSFKKFTTHGAMRKKTSS 300
QY 301 FKXSKEDDLETKKKEQAEKVDDEBEKTEPASEBOEPADDTQARLSADYEKVELPL 360
Db 301 FKXSKEDDLETKKKEQAEKVDDEBEKTEPASEBOEPADDTQARLSADYEKVELPL 360
QY 361 EDOVDDLEASSEKCAPLATEVFDEKMEAHQVVAEVAHVSTVEKTEEBEGGGGAEAGV 420
Db 361 EDOVDDLEASSEKCAPLATEVFDEKMEAHQVVAEVAHVSTVEKTEEBEGGGGAEAGV 420
QY 421 VEGTGESLPEPEKLAEPQVQAEPAEELMKSRMVCVSGDHTQULTDLSPEEKTLPKHPE 480
Db 421 VEGTGESLPEPEKLAEPQVQAEPAEELMKSRMVCVSGDHTQULTDLSPEEKTLPKHPE 480
QY 481 GIVSEVEMLSQERIKVQGSPLKPLFSSSGCLKSGKQKGRGGGGDEPREYOHINHE 540
Db 481 GIVSEVEMLSQERIKVQGSPLKPLFSSSGCLKSGKQKGRGGGGDEPREYOHINHE 540
QY 541 SPESADEOKGESSASPEPEETTCLEKPLAPODGAEEGTSDGEEKREGITPMWASF 600
Db 541 SPESADEOKGESSASPEPEETTCLEKPLAPODGAEEGTSDGEEKREGITPMWASF 600
QY 601 KKMVTPKKVRPSSESDKEELEKVKASATLSSTDSVSEMODEVKTVEGQKPEEPKRV 660
Db 601 KKMVTPKKVRPSSESDKEELEKVKASATLSSTDSVSEMODEVKTVEGQKPEEPKRV 660
QY 661 DTSVSEALICVSSKKRARKASSSDDEGCPRTLGGDSHRAEASAKDKAGCDVAVASTQ 720
Db 661 DTSVSEALICVSSKKRARKASSSDDEGCPRTLGGDSHRAEASAKDKAGCDVAVASTQ 720
QY 721 EDDOAGSSSPSPAGSPSGEGVSTWESFKRLVTPKKSGLKEEKAEDSSVEQJSTELE 780
Db 721 EDDOAGSSSPSPAGSPSGEGVSTWESFKRLVTPKKSGLKEEKAEDSSVEQJSTELE 780

```

```

QY 781 PSRESWVSIKKFIPGRKRKADKQEOATVENSQVIEINEDDPNVAVYPLSFYNAVER 840
Db 781 PSRESWVSIKKFIPGRKRKADKQEOATVENSQVIEINEDDPNVAVYPLSFYNAVER 840
QY 841 EKMAQGNTELPOLLGAVVYSEELSKTLVHTVSAVAVIDGTRAVSVREPSWISASVTE 900
Db 841 EKMAQGNTELPOLLGAVVYSEELSKTLVHTVSAVAVIDGTRAVSVREPSWISASVTE 900
QY 901 PLEHTAGAMPVEVEETKDIIAETPVLQTLPEGDAHDMDVTSVVDFTSEAVTATET 960
Db 901 PLEHTAGAMPVEVEETKDIIAETPVLQTLPEGDAHDMDVTSVVDFTSEAVTATET 960
QY 961 SEARTEVEVTAAGAEETTDVAVSAVSQLTDSPTTEATPVQVESGVLDTSEERQTA 1020
Db 961 SEARTEVEVTAAGAEETTDVAVSAVSQLTDSPTTEATPVQVESGVLDTSEERQTA 1020
QY 1021 ILQAVADKVKEESQVPAQTQVQRTGSKALEKVEEEDSEVLASEKEDVMPKGPVQAG 1080
Db 1021 ILQAVADKVKEESQVPAQTQVQRTGSKALEKVEEEDSEVLASEKEDVMPKGPVQAG 1080
QY 1081 AEHLAQSGSETQATPESLVEPEVTADVAVATCQVILQQLMEQAVAPESSETLTJDET 1140
Db 1081 AEHLAQSGSETQATPESLVEPEVTADVAVATCQVILQQLMEQAVAPESSETLTJDET 1140
QY 1141 GSTPLADSDTADGTQOQDETIDSQSKATAVROSQVTEEEAATQKEPESTLPNNVPAOE 1200
Db 1141 GSTPLADSDTADGTQOQDETIDSQSKATAVROSQVTEEEAATQKEPESTLPNNVPAOE 1200
QY 1201 EHGEPEGRDVLPTQOELTAAVAVLAKTEVQGEVQDMLDGEKVKEOEYFVHSGPNSQ 1260
Db 1201 EHGEPEGRDVLPTQOELTAAVAVLAKTEVQGEVQDMLDGEKVKEOEYFVHSGPNSQ 1260
QY 1261 KAADVTYDSEVMGAVGQCKEKSTEVQSLSEEGMETDVEKKEKTEPQVSEGEQETA 1320
Db 1261 KAADVTYDSEVMGAVGQCKEKSTEVQSLSEEGMETDVEKKEKTEPQVSEGEQETA 1320
QY 1321 APEHGTGKRVLTLDMPSSRGKALGSLGSLPPODKAGCIEVQVQSLDTVTQTALE 1380
Db 1321 APEHGTGKRVLTLDMPSSRGKALGSLGSLPPODKAGCIEVQVQSLDTVTQTALE 1380
QY 1381 AVEKVIETVVISSETGESPECVGAHLPLPAKSSATGCHMTLOHADTVPLGPGSOAESIP 1440
Db 1381 AVEKVIETVVISSETGESPECVGAHLPLPAKSSATGCHMTLOHADTVPLGPGSOAESIP 1440
QY 1441 IVPAPESLHPDLQGEISASQRESEBEDKPDAGPDAGKESSTAIEKVLKAEPEILELE 1500
Db 1441 IVPAPESLHPDLQGEISASQRESEBEDKPDAGPDAGKESSTAIEKVLKAEPEILELE 1500
QY 1501 SKSNKIVLVNVTQTVADQAPARTETAPETHAYDSQOVPAACRLDSREPRCWTMADAKKH 1560
Db 1501 SKSNKIVLVNVTQTVADQAPARTETAPETHAYDSQOVPAACRLDSREPRCWTMADAKKH 1560
QY 1561 PVQOPREDLOVLTVEAQAOPKCLPRLQKAPVSK 1596
Db 1561 PVQOPREDLOVLTVEAQAOPKCLPRLQKAPVSK 1596

RESULT 2
US-08-665-401-4
Sequence 4, Application US/08665401
GENERAL INFORMATION:
APPLICANT: Gelman, Irwin H.
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:

```

;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq Version 1.5
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/665,401
;; FILING DATE: 18-JUN-1996
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Clark, Richard S
;; REGISTRATION NUMBER: 26,154
;; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-408-2558
;; TELEFAX: 212-765-2519
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1596 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: internal
;; ORIGINAL SOURCE:
;; US-08-665-401-4

Query Match 100.0%; Score 8073; DB 10; Length 1596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAGSSTEQSPPEAGSDTSELVLSHGPAEASGAGDPADPATKLPKNGQLSS 60
DB 1 MGAGSSTEQSPPEAGSDTSELVLSHGPAEASGAGDPADPATKLPKNGQLSS 60

QY 61 VNGVAEQGDVHVQENQBGQEEVVDVQGESEDEVREKDRVEEMANSTAVEDITKDG 120
DB 61 VNGVAEQGDVHVQENQBGQEEVVDVQGESEDEVREKDRVEEMANSTAVEDITKDG 120

QY 121 QEETSEIIQIPASENNVENVQPAESQANDVGFKKVFYGFYKFTVKDKNEKSDTVQL 180
DB 121 QEETSEIIQIPASENNVENVQPAESQANDVGFKKVFYGFYKFTVKDKNEKSDTVQL 180

QY 181 LTVKDEGEAGASVAGADHQPSPVETAVGESASKESELKOSTEKQEGTLKQEOSSTEIP 240
DB 181 LTVKDEGEAGASVAGADHQPSPVETAVGESASKESELKOSTEKQEGTLKQEOSSTEIP 240

QY 241 LQAESDQAAEAEKDEGEKEKPTKSPSPSPVNSSETTSSPKFFTHGWAGRRKTS 300
DB 241 LQAESDQAAEAEKDEGEKEKPTKSPSPSPVNSSETTSSPKFFTHGWAGRRKTS 300

QY 301 FKSKEDDLETAEKKEQAEKVDDEEKEKTEPASEEQAEDTDQARLSADYKVELPL 360
DB 301 FKSKEDDLETAEKKEQAEKVDDEEKEKTEPASEEQAEDTDQARLSADYKVELPL 360

QY 361 EDQVGDLEASSEKCAPLATVFEKMAHQEVVAEAVHVSIVTEPQGGGGGAEAGGW 420
DB 361 EDQVGDLEASSEKCAPLATVFEKMAHQEVVAEAVHVSIVTEPQGGGGGAEAGGW 420

QY 421 VEGTGESLPPREKLAEPQVPOEAPAEELMKSRMCVSGGDHTQTLTDLSPREKTLPKHPE 480
DB 421 VEGTGESLPPREKLAEPQVPOEAPAEELMKSRMCVSGGDHTQTLTDLSPREKTLPKHPE 480

QY 481 GIVSEVEMLSQERIKVQGSPLKLLFSSSSGLKLSGKKQKRGKGGGDEEPEYQHIHTE 540
DB 481 GIVSEVEMLSQERIKVQGSPLKLLFSSSSGLKLSGKKQKRGKGGGDEEPEYQHIHTE 540

QY 541 SPESADEQKGESSASSPEPEETTCTLEKGPLEAPQDGAEBEGTTSDGEKKREGITPWASF 600

RESULT 3

```
US-08-978-277A-4
; Sequence 4, Application US/08978277A
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,277A
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/665,401
; FILING DATE: 18-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1596 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-978-277A-4

Query Match      100.0%; Score 8073; DB 13; Length 1596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 361 EDVQGDLEASSEKCAPLATVEPDEKMEAHQEVYAEVHVSTVETTEBEEQGGGABEGV 420
Db 361 EDVQGDLEASSEKCAPLATVEPDEKMEAHQEVYAEVHVSTVETTEBEEQGGGABEGV 420
QY 421 VEGTGESLPPEKLAEPQEVPOEAPAEELMKSRMCVSGGDHTQLTJLSPREKTLPKPE 480
Db 421 VEGTGESLPPEKLAEPQEVPOEAPAEELMKSRMCVSGGDHTQLTJLSPREKTLPKPE 480
QY 481 GIVSEVEMLSQERIKYQGSPLKKLFSSSGIKLUSKKQKGRGGGDEEPGEYOHITHTE 540
Db 481 GIVSEVEMLSQERIKYQGSPLKKLFSSSGIKLUSKKQKGRGGGDEEPGEYOHITHTE 540
QY 541 SPESADEQKSSSASPEPEETTCLEKGPLAEOODEAEFGTSSDDEKKEGTTTPMASF 600
Db 541 SPESADEQKSSSASPEPEETTCLEKGPLAEOODEAEFGTSSDDEKKEGTTTPMASF 600
QY 601 KKMVTPKKRVARPEESDKEELEKYKSATLSTSTVSEMODEVKTVGEBOKPEEPKRV 660
Db 601 KKMVTPKKRVARPEESDKEELEKYKSATLSTSTVSEMODEVKTVGEBOKPEEPKRV 660
QY 661 DTSVSWELICVGSKKRARKASSDDEGCPRTLGDSSHRAEASKDEACTDAVPASTQ 720
Db 661 DTSVSWELICVGSKKRARKASSDDEGCPRTLGDSSHRAEASKDEACTDAVPASTQ 720
QY 721 EODQAQSSSPPEPAGSPSEEGCVSTWESFKRLVTPRRKSKSLEKAKADSVEOLSTEIE 780
Db 721 EODQAQSSSPPEPAGSPSEEGCVSTWESFKRLVTPRRKSKSLEKAKADSVEOLSTEIE 780
QY 781 PSRESWSIKKFIPIGRKKRADKQEOATVEDSGPYEINDDNNVAVVPLSEYNAVER 840
Db 781 PSRESWSIKKFIPIGRKKRADKQEOATVEDSGPYEINDDNNVAVVPLSEYNAVER 840
QY 841 EKMEAQNTLPLQLGAVYSEBELSKTLVHTVSAVIDGTRAVTSVEERSPSWISASYTE 900
Db 841 EKMEAQNTLPLQLGAVYSEBELSKTLVHTVSAVIDGTRAVTSVEERSPSWISASYTE 900
QY 901 PLEHTAGAMPPEVEVETHEKQIIAETPVLTOTLPEGDAHDMDTSEVDFPSEAVTATET 960
Db 901 PLEHTAGAMPPEVEVETHEKQIIAETPVLTOTLPEGDAHDMDTSEVDFPSEAVTATET 960
QY 961 SEALTREEVTPASGAEEETDMVASVSOULTSPDTEETATPVQEVESGVLDTBEEERQQA 1020
Db 961 SEALTREEVTPASGAEEETDMVASVSOULTSPDTEETATPVQEVESGVLDTBEEERQQA 1020
QY 1021 ILQAVADKVKESQVPATQVQRTGSKALEKVEEVEDSEVLASEKEDVMPKGPVQDAG 1080
Db 1021 ILQAVADKVKESQVPATQVQRTGSKALEKVEEVEDSEVLASEKEDVMPKGPVQDAG 1080
QY 1081 AEHLAQSETCQATPESLEVEEVTADVDAVTCOVYIKLQQLMEQAVAPESSETLTDSETN 1140
Db 1081 AEHLAQSETCQATPESLEVEEVTADVDAVTCOVYIKLQQLMEQAVAPESSETLTDSETN 1140
QY 1141 GSTPLABSDTADGTQOQDETTISQDSKATAAVKQSOVTEEBEATQKKEEPTLPPNNVQOE 1200
Db 1141 GSTPLABSDTADGTQOQDETTISQDSKATAAVKQSOVTEEBEATQKKEEPTLPPNNVQOE 1200
QY 1201 EHGEPEGRDVLPTQOELTAAAVPVLAKTEVGOGEVDMLDGEKVKKEOEYFVHSGPNSQ 1260
Db 1201 EHGEPEGRDVLPTQOELTAAAVPVLAKTEVGOGEVDMLDGEKVKKEOEYFVHSGPNSQ 1260
QY 1261 KAADVTYDSEWYVAGCOEKESTEVQSLSEEGMETDVEKEKETPPEQVSEBGEQETA 1320
Db 1261 KAADVTYDSEWYVAGCOEKESTEVQSLSEEGMETDVEKEKETPPEQVSEBGEQETA 1320
QY 1321 APEHGTYGKPVLTLDMPSSRGKALSGSPSLPDDQKAGCIEVOYQSLDITVTYTOAE 1380
Db 1321 APEHGTYGKPVLTLDMPSSRGKALSGSPSLPDDQKAGCIEVOYQSLDITVTYTOAE 1380
QY 1381 AVEKVIETVVISLSTGESPEECVGAHLPLPAEKSSATGHWTLQHAEDTVPLGPESQAESLPI 1440
Db 1381 AVEKVIETVVISLSTGESPEECVGAHLPLPAEKSSATGHWTLQHAEDTVPLGPESQAESLPI 1440
```

```
Qy 1441 IVTPAPESTLHPDLQGEISASQSRSEEDKPDAGPDADGKESTAIEKVLKAEPEIILE 1500
Db 1441 IVTPAPESTLHPDLQGEISASQSRSEEDKPDAGPDADGKESTAIEKVLKAEPEIILE 1500
Qy 1501 SKSNKIVLNVITQAVDQFARTETAPETHAYDSQTVQVPACRLDSREPNCRTMKDKAKMKH 1560
Db 1501 SKSNKIVLNVITQAVDQFARTETAPETHAYDSQTVQVPACRLDSREPNCRTMKDKAKMKH 1560
Qy 1561 VVPQPREDLQVLTVLEAWAQRKCLPRQLKAPVSK 1596
Db 1561 VVPQPREDLQVLTVLEAWAQRKCLPRQLKAPVSK 1596

RESULT 4
US-09-902-432-4
; Sequence 4, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A, 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902.432
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1596
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-902-432-4

Query Match 100.0%; Score 8073; DB 23; Length 1596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGASSTEQRSPEQAGSDTSEILVLSHGHPAAASGAGPADADPATKLPOKNGQLSS 60
Db 1 MGASSTEQRSPEQAGSDTSEILVLSHGHPAAASGAGPADADPATKLPOKNGQLSS 60
Qy 61 VNGVAEQGDVHVQENQEGQEEVVDVQRESEDVREKDRVEEMANSTAVEDITKDG 120
Db 61 VNGVAEQGDVHVQENQEGQEEVVDVQRESEDVREKDRVEEMANSTAVEDITKDG 120
Qy 121 QEETSEIIIEQIPASNNVVEWQPAESQANDVGPKVFKFVGFKFTVKCKDKNEKSDTVQL 180
Db 121 QEETSEIIIEQIPASNNVVEWQPAESQANDVGPKVFKFVGFKFTVKCKDKNEKSDTVQL 180
Qy 181 LTVKKDEGEAGASVAGADHQPSPVETAVGSAKSESELKQSTKEQGTLLKQBSSTTEIP 240
Db 181 LTVKKDEGEAGASVAGADHQPSPVETAVGSAKSESELKQSTKEQGTLLKQBSSTTEIP 240
Qy 241 LQAESDQAAEAEKDEGEKEKEPTKSPSPSPVNSSETTSSFKKPTTHGAWGRKKTTS 300
Db 241 LQAESDQAAEAEKDEGEKEKEPTKSPSPSPVNSSETTSSFKKPTTHGAWGRKKTTS 300
Qy 301 FKSKEDDLETAEKRSQAEKVDDEEKEKTEPASEEQEPAEDTDQARLSADYKVELPL 360
Db 301 FKSKEDDLETAEKRSQAEKVDDEEKEKTEPASEEQEPAEDTDQARLSADYKVELPL 360
Qy 361 EDQVGDLEASSEKCAPLATVFEKMEAHQEVVAHVSTVTEKTEBERQGGGGAEGGV 420
Db 361 EDQVGDLEASSEKCAPLATVFEKMEAHQEVVAHVSTVTEKTEBERQGGGGAEGGV 420
Qy 421 VEGTGESLPPKLAEPQVPOEAPAEELMKSRMCYSGGDHTQLTDLSPBEKTLPKHPE 480
Db 421 VEGTGESLPPKLAEPQVPOEAPAEELMKSRMCYSGGDHTQLTDLSPBEKTLPKHPE 480
```

```
Qy 481 GIYSEVEMLSQERIIVQGSPLKKLFSSSGLKKLSGKKQKGRGGGDEBPGEYQHITHTE 540
Db 481 GIYSEVEMLSQERIIVQGSPLKKLFSSSGLKKLSGKKQKGRGGGDEBPGEYQHITHTE 540
Qy 541 SPESADQKSESASSPEEPETTCLEKGPLEAPQDGEAEEGTTSDEGKKEGRTTPWASF 600
Db 541 SPESADQKSESASSPEEPETTCLEKGPLEAPQDGEAEEGTTSDEGKKEGRTTPWASF 600
Qy 601 KMWTPKRVRRPESDKSEBELEKVKSATLSSDSTSVSEMQDEVKTVGEEQKPEEPKRRV 660
Db 601 KMWTPKRVRRPESDKSEBELEKVKSATLSSDSTSVSEMQDEVKTVGEEQKPEEPKRRV 660
Qy 661 DTSVSEALI CVGSSSKRKARKASSDDDEGPRTLGGDSHRAEASKOKERAGTDAVPASTQ 720
Db 661 DTSVSEALI CVGSSSKRKARKASSDDDEGPRTLGGDSHRAEASKOKERAGTDAVPASTQ 720
Qy 721 EQDQAQSSSPPEAGSPSEGEVSTWESFKRLVTPRKKSLEKKAEDSSVEQLSTEIE 780
Db 721 EQDQAQSSSPPEAGSPSEGEVSTWESFKRLVTPRKKSLEKKAEDSSVEQLSTEIE 780
Qy 781 PSREESWVSIKKFIPIGRKKRKADGKQEQATVEDSGPVEINEDDPNVPAVPLSEYNAVER 840
Db 781 PSREESWVSIKKFIPIGRKKRKADGKQEQATVEDSGPVEINEDDPNVPAVPLSEYNAVER 840
Qy 841 EKMEAQNTLPPQLLGAIVYSEBELSKTLVHTVSVAVTDGTRAVTSVEERSPSWISASVTE 900
Db 841 EKMEAQNTLPPQLLGAIVYSEBELSKTLVHTVSVAVTDGTRAVTSVEERSPSWISASVTE 900
Qy 901 PLEHTAGEAMPVVEVTEKDI IAEETPVLTQTLPEGKADHDDMVTSVDFTSEAVTATET 960
Db 901 PLEHTAGEAMPVVEVTEKDI IAEETPVLTQTLPEGKADHDDMVTSVDFTSEAVTATET 960
Qy 961 SEALRTEEVTEASGAETTTDMWSAVSOLTDSPDTTEATPVQEVESGLVDTTEEROTQA 1020
Db 961 SEALRTEEVTEASGAETTTDMWSAVSOLTDSPDTTEATPVQEVESGLVDTTEEROTQA 1020
Qy 1021 ILQAVADKVKESQVPATQVQRTGSKALEKVEBEDSEVLASEKBDKVPKGPVQEGAG 1080
Db 1021 ILQAVADKVKESQVPATQVQRTGSKALEKVEBEDSEVLASEKBDKVPKGPVQEGAG 1080
Qy 1081 AEHLAQSESTGOATPESLEVPETADVHVATCOVILKLOLMQAOVAPESSETLTDSETN 1140
Db 1081 AEHLAQSESTGOATPESLEVPETADVHVATCOVILKLOLMQAOVAPESSETLTDSETN 1140
Qy 1141 GSTPLADSDTADGTQDDTETIDSQDSKATAAVROSQVTEEEAATAQKEEPSTLPNNVPAQE 1200
Db 1141 GSTPLADSDTADGTQDDTETIDSQDSKATAAVROSQVTEEEAATAQKEEPSTLPNNVPAQE 1200
Qy 1201 EHGEPEGRDVLPTQBELTAAAVPVLAKTVEGQEGEVDWLDGKVKBEQEVFVHSGPNSQ 1260
Db 1201 EHGEPEGRDVLPTQBELTAAAVPVLAKTVEGQEGEVDWLDGKVKBEQEVFVHSGPNSQ 1260
Qy 1261 KAADVTYDSEVMGVAGCQEKESTEVQSLSEEGEMETDVEKEKETPEQVSEBGEQETA 1320
Db 1261 KAADVTYDSEVMGVAGCQEKESTEVQSLSEEGEMETDVEKEKETPEQVSEBGEQETA 1320
Qy 1321 APEHEGTGKPVLTLDMPSSERKALGSLGSPSLPDQDKAGCTIEVOVQSLDITVTOTAE 1380
Db 1321 APEHEGTGKPVLTLDMPSSERKALGSLGSPSLPDQDKAGCTIEVOVQSLDITVTOTAE 1380
Qy 1381 AVEKVIETVVISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVPPLGPESQAESIPI 1440
Db 1381 AVEKVIETVVISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVPPLGPESQAESIPI 1440
Qy 1441 IVTPAPESTLHPDLQGEISASQSRSEEDKPDAGPDADGKESTAIEKVLKAEPEIILELE 1500
Db 1441 IVTPAPESTLHPDLQGEISASQSRSEEDKPDAGPDADGKESTAIEKVLKAEPEIILELE 1500
Qy 1501 SKSNKIVLNVITQAVDQFARTETAPETHAYDSQTVQVPACRLDSREPNCRTMKDKAKMKH 1560
Db 1501 SKSNKIVLNVITQAVDQFARTETAPETHAYDSQTVQVPACRLDSREPNCRTMKDKAKMKH 1560
Qy 1561 VVPQPREDLQVLTVLEAWAQRKCLPRQLKAPVSK 1596
```

Db 1561 PVQPREDLQVLTIVLEAWAQPRLQLKAPVSK 1596

RESULT 5

PCT-US97-06830-2

Sequence 2, Application PC/TUS9706830

GENERAL INFORMATION:

APPLICANT: Gelman, Irwin H.

TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESS: Brumbaugh, Graves, Donohue & Raymond

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10112-0228

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/06830

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/665,401

FILING DATE: 18-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Richard S

REGISTRATION NUMBER: 26,154

REFERENCE/DOCKET NUMBER: A30558 - 165/34008

TELEPHONE: 212-408-2558

TELEFAX: 212-765-2519

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1346 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

PCT-US97-06830-2

Query Match 67.2% Score 5421.5; DB 1; Length 1346;

Best Local Similarity 91.6% Pred. No. 4.9e-304; Indels 15; Gaps 7;

Matches 1098; Conservative 24; Mismatches 62;

Db 387 MEAHQEVAAVAVSVTEKTEEEQGGGGEAAGGVVETGESLPPEKLAEPQEVQEAEDA 446

Db 1 MEAHQEVAAVAVSVTEKTEEEQGGGGEAAGGVVETGESLPPEKLAEPQEVQEAEDA 60

Qy 447 EELMSKREMCVSGGHTQTLTDSPEKTLPKHPEGVSVSEVMSQERKLVGGSPKTKLF 506

Db 61 EELMSKREMCVSGGHTQTLTDSPEKTLPKHPEGVSVSEVMSQERKLVGGSPKTKLF 120

Qy 507 SSSGKSLKSGKKQKRGGGGDEPGEVQHITESPESADEQKGGSSASSPPEPTTCL 566

Db 121 SSSGKSLKSGKKQKRGGGGDEPGEVQHITESPESADEQKGGSSASSPPEPTTCL 180

Qy 567 EKGPLEAPDQGEAEGETTSDGKKRE---GITPMASFKAQVTPKKYVRRPESDKEBEL 622

Db 181 EKGPLEAPDQGEAEGETTSDGKKRE---GITPMASFKAQVTPKKYVRRPESDKEBEL 234

Qy 623 EKVSATLSSDSTSVSEMODEVKTVEGEOKPEEPKRVDTSVSEALICVGSKKRARKA 682

Db 235 EKVSATLSSDSTSVSEMODEVKTVEGEOKPEEPKRVDTSVSEALICVGSKKRARKA 294

Qy 683 SSSDDEGGPRTLLGDSHRAEASDKDAGTDVAPASTQEDQAGSSSPPEAGSPSEGG 742

Db 295 SSSDIR-GPRTLLGDSHRAEASDKDAGTDVAPASTQEDQAGSSSPPEAGSPSEGG 353

Qy 743 VSTSEFRRLVTPRKKSQKLEEK-AEDSSVEQLSTELSPRESNWS--IKKIPGRK 799

Db 354 VSTSEFRRLVTPRKKSQKLEEK-AEDSSVEQLSTELSPRESNWS--IKKIPGRK 413

Qy 800 KRADKQEOATVEBDSGPEVIEDDPNPAVPLSEYNAVEREKEAOGNTLPOLLGAVY 859

Db 414 KGMGRQEOATVEBDSGPEVIEDDPNPAVPLSEYNAVEREKEAOGNTLPOLLGAVY 472

Qy 860 VSEELSKTLVHTSVAVI DGRAVTSVEBSPSWISASVTEPLEHTAGEAMPVEVTEK 919

Db 473 VSEELSKTLVHTSVAVI DGRAVTSVEBSPSWISASVTEPLEHTAGEAMPVEVTEK 532

Qy 920 DIIAETPVLTQTLPEKGDADMDVTSVDPTSEAVTATETSEALRTVEVTEASGAEET 979

Db 533 DIIAETPVLTQTLPEKGDADMDVTSVDPTSEAVTATETSEALRTVEVTEASGAEET 592

Qy 980 DMVAVSQLTDSPTTEBAPPVQEVESGLDTEEEERQTOAILQAVADKVEBQVPAQ 1039

Db 593 DMVAVSQLTDSPTTEBAPPVQEVESGLDTEEEERQTOAILQAVADKVEBQVPAQ 652

Qy 1040 TVORTGSKALEKEVEBESVLA SEKEDVMPKPYQOEAHEHIAOGSETQATPESLE 1099

Db 653 TVORTGSKALEKEVEBESVLA SEKEDVMPKPYQOEAHEHIAOGSETQATPESLE 712

Qy 1100 VPEVTADVDVATQVYIKLQQLMEQAVAPESSETLTSTETNGSTPLADSDPTAQDET 1159

Db 713 VPEVTADVDVATQVYIKLQQLMEQAVAPESSETLTSTETNGSTPLADSDPTAQDET 772

Qy 1160 IDSODSKATAVROSQVTEEBEATAQKEPSTLPNNYPAQEHGEERQVLEPTQOELT 1219

Db 773 IDSODSKATAVROSQVTEEBEATAQKEPSTLPNNYPAQEHGEERQVLEPTQOELT 832

Qy 1220 AAAPVLAKEVEGEVDMVLDGKVEBOEPVHSGPNSQKADVTYDSEVMVAGQOE 1279

Db 833 AAAPVLAKEVEGEVDMVLDGKVEBOEPVHSGPNSQKADVTYDSEVMVAGQOE 892

Qy 1280 KESTEVQSLSEBEMETDVAKETKPEBOVSSEGOETAPHEGTYGKPVYTLMP 1339

Db 893 KESTEVQSLSEBEMETDVAKETKPEBOVSSEGOETAPHEGTYGKPVYTLMP 952

Qy 1340 SERGKALGSLGSPSLPDQDAGCIEVQVOSLDTTVOATAVEKVIETVVISSEGESPE 1399

Db 953 SERGKALGSLGSPSLPDQDAGCIEVQVOSLDTTVOATAVEKVIETVVISSEGESPE 1012

Qy 1400 CVGAHLPAEKSSATGGMWTLQHAEDTVPLGESQAESIPITVPAPBSTLHPDLOGEIS 1459

Db 1013 CVGAHLPAEKSSATGGMWTLQHAEDTVPLGESQAESIPITVPAPBSTLHPDLOGEIS 1072

Qy 1460 ASQRESEEBEKPDPAGPADGKESTALIEKYLKAPPELIELESKSNKYLVNIQTAVDQFA 1519

Db 1073 ASQRESEEBEKPDPAGPADGKESTALIEKYLKAPPELIELESKSNKYLVNIQTAVDQFA 1132

Qy 1500 RTEPAETHAVDSOTQVPAACGLDSREPNRCWTMMKDKMKHPVQOPREDLVLTIVLEAW 1578

Db 1133 RTEPAETHAVDSOTQVPAACGLDSREPNRCWTMMKDKMKHPVQOPREDLVLTIVLEAW 1191

RESULT 6

US-08-665-401-2

Sequence 2, Application US/08665401

GENERAL INFORMATION:

APPLICANT: Gelman, Irwin H.

TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESS: Brumbaugh, Graves, Donohue & Raymond

STREET: 30 Rockefeller Plaza

CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10112-0228
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/665,401
 FILING DATE: 18-JUN-1996
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Richard S
 REGISTRATION NUMBER: 26,154
 REFERENCE/DOCKET NUMBER: A30558 - 165/34008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-408-2558
 TELEFAX: 212-765-2519
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1346 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: internal
 ORIGINAL SOURCE:
 US-08-665-401-2

Query Match 67.2%; Score 5421.5; DB 10; Length 1346;
 Best Local Similarity 91.6%; Pred. No. 4.9e-304;
 Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;
 QY 387 MEAHEVVAEVHVTVEKTEEEQGGGGAEGGVVVEGTGSLPEKLAEPQVEPQEA 446
 DB 1 MEAHEVVAEVHVTVEKTEEEQGGGGAEGGVVVEGTGSLPEKLAEPQVEPQEA 60
 QY 447 BELMKSREMCVSGGDHTQLTDLSPKPEKTLPHKPEGIIVSEVEMLSQRIKVGQSLKLF 506
 DB 61 BELMKSREMCVSGGDHTQLTDLSPKPEKTLPHKPEGIIVSEVEMLSQRIKVGQSLKLF 120
 QY 507 SSSGLKSLGKKQKRGGGDEEPGYQHHTSPESADEQKGESSASGPEEPTTCL 566
 DB 121 SSSGLKSLGKKQKRGGGDEEPGYQHHTSPESADEQKGESSASGPEEPTTCL 180
 QY 567 EKGPLEAPOGEAEGTTSDEKKEKRE----GITWAFKMTVPKRVRRPSESDEKEEL 622
 DB 181 EKGPLEAPRMGKLKELLR-GEKKRKHSLGI-----LQKDGDTQETVRRPSESDEKEEL 234
 QY 623 EKVSATLSSTDSIVSEMDEKVTGVEOKPEEPKRRVDTSVSWEALICVSSSKRKARKA 682
 DB 235 EKVSATLSSTDSIVSEMDEKVTGVEOKPEEPKRRVDTSVSWEALICVSSSKRKARKA 294
 QY 683 SSSDDEGGPRTLGGDSHRAEASKDXEAGTDAVPASTQEQDAQSSSPAGSPSEGG 742
 DB 295 SSSDIR-GPRTLGGGQSQRGQQRSRDTAVPASTQEQDAQSSSPAGSPSEGG 353
 QY 743 VSTWESFKRLVTPPKSKSKLEEK-ADDSVEQLSTIEPSPRESWVS--IKKFIPOGRK 799
 DB 354 VSTWESFKRLVTPPKSKSKLEEKAGRTLLVVGAGCVRNRVKNLGFPLRNSPDGGR 413
 QY 800 KRADKQEQATVEDSGPVEINEDDPNPVAVVPLSEYNAREKMEAGNTELPQLLGAVY 859
 DB 414 KGQMGREQATVEDSGPVEINEDDPNPVAVVPLSEYNAREKMEAGNTELPQLLGAVY 472

QY 860 VSELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEK 919
 DB 473 VSELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEK 532
 QY 920 DIIAETPVLVTQTLPEKDAHDDMTSEVDFTSEAVTATSETSEALRTEETVEASGABETT 979
 DB 533 DIIAETPVLVTQTLPEKDAHDDMTSEVDFTSEAVTATSETSEALRTEETVEASGABETT 592
 QY 980 DMVSQVQLTDSPTTTEATPVQEVESGVLDTBEERQTQAILQAVADKVKESQVPA 1039
 DB 593 DMVSQVQLTDSPTTTEATPVQEVESGVLDTBEERQTQAILQAVADKVKESQVPA 652
 QY 1040 TVQRTGSKALEKVEEVEEDSEVLASEKEKDMPKGPVQVQVQVQVQVQVQVQVQV 1099
 DB 653 TVQRTGSKALEKVEEVEEDSEVLASEKEKDMPKGPVQVQVQVQVQVQVQVQVQV 712
 QY 1100 VPEVTADVDHVATCOVIKLOLMQVAVPESSETLDTSETNGSTPLADSDTADTQODET 1159
 DB 713 VPEVTADVDHVATCOVIKLOLMQVAVPESSETLDTSETNGSTPLADSDTADTQODET 772
 QY 1160 IDSQDSKATAAQRQSVQVTEEEAATAQKEEPTSLPNNVPAQEEHGEERGRDVLPTQOELT 1219
 DB 773 IDSQDSKATAAQRQSVQVTEEEAATAQKEEPTSLPNNVPAQEEHGEERGRDVLPTQOELA 832
 QY 1220 AAAPVLAKTEVQEGEVVDWLDGKVKKEQEVFVHSGPNSQKAADVTYDSEVMGVACQ 1279
 DB 833 AAAPVLAKTEVQEGEVVDWLDGKVKKEQEVFVHSGPNSQKAADVTYDSEVMGVACQ 892
 QY 1280 KESTEVQSLSEEGEMETDVEKRETKPKQVSEEGEQETAAPHEGTYCKPVLTLDMPS 1339
 DB 893 KESTEVQSLSEEGEMETDVEKRETKPKQVSEEGEQETAAPHEGTYCKPVLTLDMPS 952
 QY 1340 SERGKALGSLGSPSLPDQDKAGCIEVQVQSLDTTQTAEAVEKVIETVVISSETGESPE 1399
 DB 953 SERGKALGSLGSPSLPDQDKAGCIEVQVQSLDTTQTAEAVEKVIETVVISSETGESPE 1012
 QY 1400 CVGAHLLPAEKSSATGSHWTLQHAEDTVPLGSPQSAESPIIIVTPAPESTLHPDLQGEIS 1459
 DB 1013 CVGAHLLPAEKSSATGSHWTLQHAEDTVPLGSPQSAESPIIIVTPAPESTLHPDLQGEIS 1072
 QY 1460 ASORESEEDKDPADGDKESTALEKVLKAEPEILELESKSNKIVLNVITAVDQFA 1519
 DB 1073 ASORESEEDKDPADGDKESTALEKVLKAEPEILELESKSNKIVLNVITAVDQFA 1132
 QY 1520 RTETAPETHAYDSQTVPAACRLDSREPNCWTWKDAKMKHPVQPPREDLQVLTVLEAW 1578
 DB 1133 RTETAPETHAYDSQTVPAACRLDSREPNCWTWKDAKMKHPVQPPREDLQVLTVLEAW 1191
 RESULT 7
 US-08-978-277A-2
 ; Sequence 2, Application US/08978277A
 ; GENERAL INFORMATION:
 ; APPLICANT: Gelman, Irwin H.
 ; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
 ; STREET: 30 Rockefeller Plaza
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10112-0228
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978,277A
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1346 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-978-277A-2

Query Match      67.2%; Score 5421.5; DB 13; Length 1346;
Best Local Similarity 91.6%; Pred. No. 4.9e-304;
Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;

QY 387 MEAHQEVAAVHVSTVETKTEEEQGGGAEAGGVVETGESLPPEKLAEPQVEQAEBA 446
DB 1 MEAHQEVAAVHVSTVETKTEEEQGGGAEAGGVVETGESLPPEKLAEPQVEQAEBA 60
QY 447 BELMKSREMCVSGGHTQTLTLSPPEKTLPKHPEGIIVSEVEMLSQERIKVQSGSLKLF 506
DB 61 BELMKSREMCVSGGHTQTLTLSPPEKTLPKHPEGIIVSEVEMLSQERIKVQSGSLKLF 120
QY 507 SSSGKXKLSGKKQKRGKGDEEPGEYOHHTESPESADQKSGSSASPEPEETTC 566
DB 121 SSSGKXKLSGKKQKRGKGDEEPGEYOHHTESPESADQKSGSSASPEPEETTC 180
QY 567 EKGPLEAPQDGEAEGETTSDGKKRE---GITPMASFQKVTPKKRVRRPESDKEBEL 622
DB 181 EKGPLEAPQDGEAEGETTSDGKKRE---GITPMASFQKVTPKKRVRRPESDKEBEL 234
QY 623 EKVKSATLSTSTSVSEMODEVYKTVGEBOKPEEPKRVDTSVSWELICVGSKKRARA 682
DB 235 EKVKSATLSTSTSVSEMODEVYKTVGEBOKPEEPKRVDTSVSWELICVGSKKRARA 294
QY 683 SSSDDEGGRTLGSDSHRAEASKQKEAGTDAVPASTOEDQAQSSSPPEPAGSPSEGE 742
DB 295 SSSDIR-GEITLGGGOSGRGQORRRTDAVPASTOEDQAQSSSPPEPAGSPSEGE 353
QY 743 VSTWSEFKRLVTPRKKSXSKLEEK-AEDSSVEQLSTELSPSSEGSVW--IKKFIPIRRK 799
DB 354 VSTWSEFKRLVTPRKKSXSKLEEKAGRTLVVAGCPLNSVVEKVLGPLNNSPDGR 413
QY 800 KRAQKQKQATVEDSGPVEINEDDVPVAVVPLSEINAVEREKMEKQGTLPOLLGANY 859
DB 414 KQOMQROEQATVEDSGPVEINEDDVPVAVVPLSEINAVEREKMEKQGTLPOLLGANY 472
QY 860 VSEELSKTLVHVSAVAVIDGTAAVTSVEERSPSWISASVTEPLEHTAGAMPVEEVTEK 919
DB 473 VSEELSKTLVHVSAVAVIDGTAAVTSVEERSPSWISASVTEPLEHTAGAMPVEEVTEK 532
QY 920 DIIAETPVLITLPEKGAHDMVTSEVDTSEAVTATETSEALRTEEVTEASGAETT 979
DB 533 DIIAETPVLITLPEKGAHDMVTSEVDTSEAVTATETSEALRTEEVTEASGAETT 592
QY 980 DMVASVQSITDSDPTTEATPVQVEESGVLDTBEERQTOAILQAVADVKESQVPAQ 1039
DB 593 DMVASVQSITDSDPTTEATPVQVEESGVLDTBEERQTOAILQAVADVKESQVPAQ 652
QY 1040 TYQRTGSKALEKEVEEEDSEVLASEKEDVMPKGPVQEAAGAEHLAQSGSETQATPESLE 1099

```

```

DB 653 TYQRTGSKALEKEVEEEDSEVLASEKEDVMPKGPVQEAAGAEHLAQSGSETQATPESLE 712
QY 1100 VPEVTADVHVATQVITKLQOLMEQAVAPSSSETLTSTETNGSTPLADSDTADTQDET 1159
DB 713 VPEVTADVHVATQVITKLQOLMEQAVAPSSSETLTSTETNGSTPLADSDTADTQDET 772
QY 1160 IDSQSKATAAVROSQVTEEEAATAQKEPSTLPPNNPAQOEHEEGEGRDPLDPTQOELT 1219
DB 773 IDSQSKATAAVROSQVTEEEAATAQKEPSTLPPNNPAQOEHEEGEGRDPLDPTQOELT 832
QY 1220 AAAPVPLAKTEVGEQGEVDMLDGEKVEEDEVFVHSGPNSQKADVTYDSEVMVAGCOE 1279
DB 833 AAAPVPLAKTEVGEQGEVDMLDGEKVEEDEVFVHSGPNSQKADVTYDSEVMVAGCOE 892
QY 1280 KESTEVQSLSEEBEMETDVEKRETKPBOVSSEGEOTPAPEHETGYKPVLTLMPS 1339
DB 893 KESTEVQSLSEEBEMETDVEKRETKPBOVSSEGEOTPAPEHETGYKPVLTLMPS 952
QY 1340 SERGKALGSLGSPSLPDQDQKAGCIEVQVOSLDTTVOATAEAVKVIETVVISETGESPE 1399
DB 953 SERGKALGSLGSPSLPDQDQKAGCIEVQVOSLDTTVOATAEAVKVIETVVISETGESPE 1012
QY 1400 CVGAHLPLAEKSSATGGMWTLQHAEDTVPLGPEQASIPITVPAPESTLHPDLQGEIS 1459
DB 1013 CVGAHLPLAEKSSATGGMWTLQHAEDTVPLGPEQASIPITVPAPESTLHPDLQGEIS 1072
QY 1460 ASQERSEEDKPDGADGKESTALIEKVLKAPPELLELSKSNKVLVNIQTAVDOFA 1519
DB 1073 ASQERSEEDKPDGADGKESTALIEKVLKAPPELLELSKSNKVLVNIQTAVDOFA 1132
QY 1520 RTEPAETHAVDSQTOVPACGLDSREPNCWTMMKDKMKHPVQOPREDLOVLTVLEAM 1578
DB 1133 RTEPAETHAVDSQTOVPACGLDSREPNCWTMMKDKMKHPVQOPREDLOVLTVLEAM 1191

RESULT 8
US-09-902-432-2
; Sequence 2, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156, 0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1346
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-902-432-2

Query Match      67.2%; Score 5421.5; DB 23; Length 1346;
Best Local Similarity 91.6%; Pred. No. 4.9e-304;
Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;

QY 387 MEAHQEVAAVHVSTVETKTEEEQGGGAEAGGVVETGESLPPEKLAEPQVEQAEBA 446
DB 1 MEAHQEVAAVHVSTVETKTEEEQGGGAEAGGVVETGESLPPEKLAEPQVEQAEBA 60
QY 447 BELMKSREMCVSGGHTQTLTLSPPEKTLPKHPEGIIVSEVEMLSQERIKVQSGSLKLF 506
DB 61 BELMKSREMCVSGGHTQTLTLSPPEKTLPKHPEGIIVSEVEMLSQERIKVQSGSLKLF 120
QY 507 SSSGKXKLSGKKQKRGKGDEEPGEYOHHTESPESADQKSGSSASPEPEETTC 566

```



```
Db 121 SSSGLKLLSGKKKGKGGGDEBPGEYQVHIHTESPESADEQKGESSASSPERPEETTCL 180
Qy 567 EKGLEAPQGEABEGTTSDEKKRE-----GITWASFKKMVTPKKVRPSPESDKKEEL 622
Db 181 EKGLEAPRNGKURKELLR-GEKKDHSLG1-----LQKDGDTQETVRRPSPESDKKEEL 234
Qy 623 EKVSATLSSTDSVTSWQDEVTGVEQKPEEPKRRVDTSVSWEALICVSSSKKRARKA 682
Db 235 EKVSATLSSTDSVTSWQDEVTGVEQKPEEPKRRVDTSVSWEALICVSSSKKRARKA 294
Qy 683 SSSDDEGPRTLGGDHSRAEASAKDEAGTDVAVPASTQEQDQAQSSSPBPAGSPSSEGG 742
Db 295 SSSDIR-GRPTLGGGQSQRGQRSRTRDVPASTQEQDQAQSSSPBPAGSPSSEGG 353
Qy 743 VSTWESFKRLVTPRKKSLEEK-ADSSVEQLSTIEBPSRBSWVS--IKFIPORRK 799
Db 354 VSTWESFKRLVTPRKKSLEEKAGRTLGVGAGCFLRSNRVEKNLGFPLRNSPPDGR 413
Qy 800 KRADGKQEQATVEDSGPVEINEDPNVPVAVPLSEYNAREKMEAGNTELPQLLCVY 859
Db 414 KQMGREQATVEDSGPVEINEDPDVPVAVPLSEYDAVEREKMEAGNTELPSCWGCV- 472
Qy 860 VSELSKTLVHTSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEEYTEK 919
Db 473 VSELSKTLVHTSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEEYTEK 532
Qy 920 DIIAETPVLVTLPEKGKADHDMVTSEVDTSEAVTATITSEALRTEEVTEASGABETT 979
Db 533 DIIAETPVLVTLPEKGKADHDMVTSEVDTSEAVTATITSEALRTEEVTEASGABETT 592
Qy 980 DMVSASQLTDSPTTTEATPVQVEGVLDTREEROTQAILQAVADVKKESQVPATQ 1039
Db 593 DMVSASQLTDSPTTTEATPVQVEGVLDTREEROTQAILQAVADVKKESQVPATQ 652
Qy 1040 TVORTGSKALEKVEVEEDEDSEVLASEKEKVMKPGPVQEAHAELAGCSETGQATPSLE 1099
Db 653 TVORTGSKALEKVEVEEDEDSEVLASEKEKVMKPGPVQEAHAELAGCSETGQATPSLE 712
Qy 1100 VPEVTADVHVATCQVTKLQOLMEQAVAPESSETLTDSNTGSLPLADSDTADTQODET 1159
Db 713 VPEVTADVHVATCQVTKLQOLMEQAVAPESSETLTDSNTGSLPLADSDTADTQODET 772
Qy 1160 IDSODSKATAVRQSQVTEBEAATAQKEEPTLBNVPQAQEEGEEPRGRDVLPTQOELT 1219
Db 773 IDSODSKATAVRQSQVTEBEAATAQKEEPTLBNVPQAQEEGEEPRGRDVLPTQOELA 832
Qy 1220 AAAPVPLAKTEVGQEGVDMLDGKVKKEBOEVFVHSGPNQKAAADVTDSEVMGVACQE 1279
Db 833 AAAPVPMQKTEVGQEGVDMLDGKVKKEBOEVFVHSGPNQKAAADVTDSEVMGVACQE 892
Qy 1280 KESTEVQSLSEEGEMETDVEKRETKPEQVSEGEQETAAPHEGTYGKPVLTLDMP 1339
Db 893 KESTEVQSLSEEGEMETDVEKRETKPEQVSEGEQETAAPHEHRYGKPVLTLDMP 952
Qy 1340 SERCKALGSLGGSPSLPDQKAGCIEVQVQSLDTTDTTQTAABAVKVTETVVISETGSPE 1399
Db 953 SERCKALGSLGGSPSLPDQKAGCIEVQVQSLDTTDTTQTAABAVKVTETVVISETGSPE 1012
Qy 1400 CVGAHLLPAEKSSATGHTWLQHAEDTVPLGPESQAESIPITVPAPESTLHPDQGEIS 1459
Db 1013 CVGAHLLPAEKSSATGHTWLQHAEDTVPLGPESQAESIPITVPAPESTLHPDQGEIS 1072
Qy 1460 ASQRSEEDKPDAGPADGKESTAIEKVLKAEPEILELESKNKVLNVIQTAVDQFA 1519
Db 1073 ASQRSEEDKPDAGPADGKESTAIDKVLKAEPEILELESKNKVLNVIQTAVDQFA 1132
Qy 1520 RTETAPETHAYDSQTPACRLDSREPNRCWTMKDAMKHPVPQPPREDLOVLTVEAW 1578
Db 1133 RTETAPETHAYDSQTPAMRLDSREPNRCWTMKAMKHPVPQPPREDLOVLTVEAW 1191
```

RESULT 9

US-60-389-987-2097

```
; Sequence 2097, Application US/60389987
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465P2
; CURRENT APPLICATION NUMBER: US/60/389,987
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2097
; LENGTH: 1782
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-389-987-2097
```

```
Query Match 53.5%; Score 4318.5; DB 27; Length 1782;
Best Local Similarity 58.1%; Pred. No. 4.7e-240;
Matches 961; Conservative 189; Mismatches 398; Indels 107; Gaps 34;
```

```
Qy 1 MGAGSSTEQRSPQ-PAGSDTPSELVLSGHGPAAEAS-GAAGDPADA--DPATKLPQKNG 56
Db 1 MGAGSSTEQRSPQPEGSSSTPAEPSPGGPSAAAPDTPADPAIAASDPATKLLQKNG 60
Qy 57 QLSSVNGVAEGDVHVQENQEQ-----EEVVDEVDQGRESDEVREKDR 102
Db 61 QLSTINGVAQDELSQEGDLNGKALNGQAGALNQEEVEVITVEVGQSDSDSVKRD 120
Qy 103 VEEMANSTAVEDITKQGESETSEITQIPASNNVVEVMQPAESQANDVGPKKVPKFVG 162
Db 121 DKEMATKSAVVDITDDQGEETPEITQIPSSSNLEELTQPTESQANDLGFKKVPKFVG 180
Qy 163 FKFTVKDKNEKSDTVQLLTIVKKDEGEAEASVAGADHQPSPSVETAVGSESASELSKQS 222
Db 181 FKFTVKDKTEKPDVTQLLTIVKKDEGEA--AGAGDHKDPDL--GAGEAASESEPKQS 235
Qy 223 TEKQEGTLKQEQSSTPIQAESDQAAEBEAKDGEKEKEPTKSPSPSPVNETTS 282
Db 236 TEKPEETLKREQSHAEISPPAESQAV-BECKEKEKEKEKEPSKSAESPTSPVTSETGS 294
Qy 283 SPKKFTTHGWAGWRKTSFKKSKEDDLETAEKKEQEAKEVDEEKEKTEPASEE---- 337
Db 295 TFKKFTQGWAGWRKTSFKPKEDDEVEASEKKKEQEPKVDTEEDGKAEVASEKLTASE 354
Qy 338 ----QBPEDTDQARLSADYKVELPLEDQVGDLEASSBEKCAPLATEVDFDEKWEAHQ-E 392
Db 355 QAHPQEPASAEHPRLSAEYKVELPSEBQVSGSQSPSEKAPLATEVDFDEKLVHQEE 414
Qy 393 VVAEVHVTVEKTEBQGGGGAEGGVVVEGTGESLPPKLAEPQVQEPQEAEPABELMKS 452
Db 415 VVAEVHVTVEERTEQ-----RTEVEETAGSVPASELVEMDAPQEAEPASLVL 466
Qy 453 REMCVSGDHTQLTDLSPBEKTLPKHPEGIVSEVEMLSQERIKVQGSPLKKLFPSSGLK 512
Db 467 KETCVSGEDPTQAGADLSKPEKVLSPPEGVVSEVEMLSQERMKVQGSPLKFLTSTGLK 526
Qy 513 KLSGKKQKRGKGGGDEEPGEYQVHIHTESPESADEQKGESSASSPEPETTTCLEKGP 572
Db 527 KLSGKKQKGR--GGGDESEGEHTQVPADSPDSQEEQKGESSASSPEPEITCLEKGLAE 595
Qy 573 APODGABEGTTSDEKKKEGITPWASFKKMVTPKKVRPSPESDKKEELEKVKATLSS 632
Db 586 VQDGEABEGATSDGKKEGVTWASFKKMVTPKKVRPSPESDKEDBLDKVKSATLSS 645
Qy 633 TDSVSEMQDEVKTVGEEQKPEPKRRVDTSVSWEALICVSSSKKARKASSSDDEGGPR 692
Db 646 TESTASEMQEEMKGSVEEPKPKPKRVDTSVSWEALICVSSSKKARKASSSDDEGGPK 705
```

```

QY 693 TLGDSHRAEASKDKKAGTDAVPASTQDQDQAGSSSPREPAGSPSEGBVSTWMSFKRL 752
Db 706 AMGSHQKADKAGKDKETGTDGIIAGSQHDPGQSSSSPEQSPTEGBGVSTWMSFKRL 765
QY 753 VTPRKSKSKLEEKAD-----SSVEQLSTEIPRSREESVSIKKFIPIGRKKRAAGKQBO 808
Db 766 VTPRKSKSKLEEKEDSISAGSGVHSHSTPDTIPGKEBSVSIKKFIPIGRKKRPPGKQBO 825
QY 809 ATVEDSGVEINEDDPNPAVAVPLSEYNAREKME---AQNTLPLOLLGAVVYSELS 865
Db 826 APVEDAGETGANNEDSDPAVAVPLSEYDAVEREKMEAOQAQASAPQEPKATVSKELS 885
QY 866 KTLVHTSVAVVIDGTAVTSVEERSPSWISASVTEPLEHTAGEAMPVHEVTEKIIA-E 924
Db 886 ESQVHMMMAAAVADGTAAATIIIEERSPSWISASVTEPLEBEVEAALLTBEVEREVIABE 945
QY 925 ETPVLTOTLPEKADHDMVNTSEVDPTSEAVATETSEALRTVEVTEASGAETDWMVA 984
Db 946 EPTVTEPLEPENREARQDVTVSEAEITTPRAVTAETAGPLGAEGETEASAEETTEMVA 1005
QY 965 VSQLTDSPTTEATPVQEVESGVLDTEEEERQTAIIQAVADKVKESQVPAT----- 1038
Db 1006 VSQLTDSPTTEATPVQEVESGVLDTEEEERQTAIIQAVADKVKESQVPAT----- 1065
QY 1039 -QTVQRTGSKALEKVBVEVEDESEVLASEKQVMKGPVQEGAEHLAQSGSETGQATPES 1097
Db 1066 LQPVQRA-----EAEPRPEQAEASGLKKEKTVVLKVDQAEKTEPTFGKVVQQTTPES 1119
QY 1098 LE-VREVTAADVH---VATQCV-----IKLOOL-MEQAVAPSSSETLDTSETNGSTPLAD 1147
Db 1120 FKAQVQVESITSSSELVITQAEETLAGVKSQEMWEOALPPSVETPTTSETDGSFPVAD 1179
QY 1148 SPTADGTOQDETITDSQSKATAVAVQSVTEBEAATQKEEPTLNNVPAQEHGEERG 1207
Db 1180 FPARPTQODELVEIHENBVA SGTSQGTAEAVPAQERPA-PSSVFPQETKEQSK 1238
QY 1208 -RDVLEPTQOELTAAVAVPLATEVQEGEVDMLDGEKXK-----EQQEVFVHSG--PMS 1259
Db 1239 MEDTLEHTDKEVSVEVTSIISTKEGTQ--EADQYADEKTKDVPFPGLEGSIDTGIIVSR 1296
QY 1260 QKAADVTVSEVMGVAGCOEKSTEVOS-----LSLEGEEMTDVYKEREKTEPQEVSE 1314
Db 1297 EKVTEVAKGEGSTEVECKKDALTELQSHAKSPSPSEVEEMVAVYEREKTEAPHTVNEE 1356
QY 1315 G-EOETAAHEHGTGKPLVLTLDMPSSERKALGSLGS--PSLPDQKAGCIEVQVOSLD 1372
Db 1357 KLEHETAVVSEBVSQQLQTVNVPIIDAKAVSLESGSPPLQGGAEAVCTKIQVQSSB 1416
QY 1373 TTVTQTAAEVAKV--ETVVISETGSPSCVGAHLLPAKSSATGHWTLQHAEDTVPLG 1430
Db 1417 ASFTLTAAAEKKEVLEGTANIILETGETLTPAGAHVLEKSSSEKNEFDAHGEEDAVPTG 1476
QY 1431 PESQAESITITTPABESTLHPLOGEIASQERESEEDKPDAGDAGKSTALEKXL 1490
Db 1477 PQCQASTVIVSATTKGGLSSDLSEKETSILKMKSDVEDEVOVACQEV--KISVALIEDU- 1533
QY 1491 KAEPE--IILELSSKNKIULANVQITAVDOFART-ETAPETHAVDSQTOVPACRLDSREBN 1547
Db 1534 --EPENGIIIELETKSKVLQNIITQITAVDOFVTRTEFATMLSELTOQHAVIKADSQDAG 1591
QY 1548 KCMWK-----MKDAKMHFVPOPREDLQVLTVLEA 1577
Db 1592 QETKEGEPQASQODETPTTSKESSESTAVGQA 1626

```

RESULT 10

```

; Sequence 2097, Application US/60412418
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fany, Bojin D.
; APPLICANT: Zhang, Bing

```

```

; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465P3
; CURRENT APPLICATION NUMBER: US/60/412.418
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2097
; LENGTH: 1782
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-412-2097

Query Match      53.5%; Score 4318.5; DB 27; Length 1782;
Best Local Similarity 58.1%; Pred. No. 4.7e-240;
Matches 961; Conservative 189; Mismatches 398; Indels 107; Gaps 34;

QY 1 MGAGSTTEQRSPEQ-PAGSDTPESELVLSGHCPAAEAS-GAAGDPADA--DPATYLPQKNG 56
Db 1 MGAGSTTEQRSPEQPEGSSTPAEPSPGGGSAEAPDTTADAPAIASDPATYLPQKNG 60
QY 57 QLSNVGVAEGDVHVGNEQO-----EEEVNEDVQORESEVREKDR 102
Db 61 QLSNVGVAEGDELISLGDNLQKALNGQALNSQEEBEVITVEGQRISEVSKDS 120
QY 103 VEEMANSTAVEDIITKQOEEETSEIIEQIPASENNVEEMVOPAESQANDVFPKTVFKVG 162
Db 121 DKMATKSAVVDITDQGEETPEIIRIPISSSENLBELTPTESQANDIQFKVFKVG 180
QY 163 FKFTYKDKNEKSDTVQLLITVKQDGBGAAEAVGAGDHQEPSTVAVGESAKSELSKOS 222
Db 181 FKFTYKDKNEKSDTVQLLITVKQDGBGAAEAVGAGDHQEPSTVAVGESAKSELSKOS 235
QY 223 TEKQGTLKQOSSTEIPLQAESDQAAEEAKDGESEKQEKPEKSPSPSPVNSSETTS 282
Db 236 TEKPEETIKRQSHAEISPPAESQAV-EECKEKEEKEKPEKSPSPSPVNSSETTS 294
QY 283 SPKKEFTGMAKMRKTKTSPKSKKEDLETAERKKEQAEKVEDEKTEPASEE----- 337
Db 295 TYKKEFTGMAKMRKTKTSPKSKKEDLETAERKKEQAEKVEDEKTEPASEE----- 354
QY 338 ---QEPREDTQARLSADYKVELPLEDQYDGLFASSESEKCAPLATEVPEKMEAHQ-E 392
Db 355 QAHPEPAPESAHBRLSAEYKVELPSEBQVSGSGPSEKAPLATEVPEKMEAHQ-E 414
QY 393 VVAEYHVAETVEKTEBEGQGGGABGQVVEGTGESLPPEKLAEPQEVQAEAPAEMLKMS 452
Db 415 VVAEYHVAETVEKTEBEGQGGGABGQVVEGTGESLPPEKLAEPQEVQAEAPAEMLKMS 466
QY 453 REMCVSGGDHQLTLDLSPKEKTLPHKPEGIYSEVEMLSQERIKYQSGPLKLFPSSGLX 512
Db 467 KETCVSGEDPLQAGDLSDPEKVLSPPEGVSEVEMLSQERIKYQSGPLKLFPSTGLX 526
QY 513 KLSGKKQKGGGGGDEEPGEVQHITHSPESADQKGESSASPSPEDEFTTCLKGPLE 572
Db 527 KLSGKKQKGRK-GGDEBSGHEHTQVPADSPSOERQKGSASPSPEDEFTTCLKGLAE 585
QY 573 APQDEABEGTTSODEKRGREGITTPWASFKXVNTPKKRRRRSESEKPELEKXVATLSS 632
Db 586 VOQDEABEGGASDDEKREKREGVTPWASFKXVNTPKKRRRRSESEKPELEKXVATLSS 645
QY 633 TDSVSEMODEVKTIVGEEQKPEEPKRRVDTSVSWALICVSSSKKRRARASSDDEGGPR 692
Db 646 TDSVSEMODEVKTIVGEEQKPEEPKRRVDTSVSWALICVSSSKKRRARASSDDEGGPR 705
QY 693 TLGDSHRAEASKDKKAGTDAVPASTQDQDQAGSSSPREPAGSPSEGBVSTWMSFKRL 752
Db 706 AMGSHQKADKAGKDKETGTDGIIAGSQHDPGQSSSSPEQSPTEGBGVSTWMSFKRL 765

```

```
Qy 753 VTPRKSKSLEEKAE-----SSVEQLSTETEPSREESWVSIKKFIPGRRKKRADGKQEQ 808
Db 766 VTPRKSKSLEEKSEDSIAGSGVHSTPDTEPGKSESWVSIKKFIPGRRKKRPDQKQEQ 825
Qy 809 ATVEDSGVINEDDPNVPAVPLSEYNAREKME---ACNTELPOLLGAVVYSBELS 865
Db 826 APVEDAGPTGANEDSDVPAVPLSEYDAVERKWEAQAKSAEQEPQAKATEVSKELS 885
Qy 866 KTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPVVEVTEKDI1A-E 924
Db 886 ESQVHMAAAVADGTTRAATIIERSPSWISASVTEPLEQVEAEAAALLTEEVLEEVIAEE 945
Qy 925 ETPVLTOTLPGKADHDMVTSEVDFTSEATATSETSEALTEETVTEASGAETTDVUSA 984
Db 946 EPTVTEPLPENRGARTVVSEAELETPAVTAETAGPLGAEGBTEASAAEETEMVUSA 1005
Qy 985 VSQLTSDPTTEETATPVQEVESGLVDTEEBERQTQAILOAVADVKKEESQVPAT----- 1038
Db 1006 VSQLTSDPTTEETATPVQEVESGLVDTEEBERQTQAILOAVADVKKEESQVPAT----- 1065
Qy 1039 -QTVORTGSKALEKVEEVEEDSEVLASEKEKDVMPKGPVQBGAGAEHLAQGSETQOATPES 1097
Db 1066 LQPVQRA-----BAERPEEQAEASGLKEDTVLVKVDQAQEAKEPFTQGVVGTTPES 1119
Qy 1098 LE-VPEVTADVH---VATCOV-----IKLOOL-MEQAVAPESSETLTDSETNGSTPLAD 1147
Db 1120 FEKAPQVTEESSELVTTQOAEFLAGVQSEMYMEQAIPDPSVETPTDSETDGTSTPAD 1179
Qy 1148 SDAAGTQODTIDSQSKATAAARQSOVTEEEAATAQKEPSTLPNNVPAQEGHREPG 1207
Db 1180 FDAQTTQKDIIVEIHENEVASGTQSGGTEAEAVPAQKERRPPA-PGSFVQEETKQSK 1238
Qy 1208 -RDVLEPTQBELTAAAVPVLAKTEVGQBEVDWLDGEKV-----EEOQVVFVHSG--PNS 1259
Db 1239 MEDTLEHTKEVSVETVSIKSTEGTQ--EADQVADKTKDVPPFEGLEGSIDGTIVSR 1296
Qy 1260 QKAADVITYDSEVMGACQEKESTEVQS-----LSLEGEMETDVEKRETKTEQVSEEE 1314
Db 1297 EKVTEVALKGBGTBEAECKODDALELOSHAKSPSPVEREMVQVVEREKTAEAEPTHYNEE 1356
Qy 1315 G-EQETAAPHEGTVGKPVLTLDMPSSERKALGSLGGS-FSLPDQDKAGCIEVQVOSLD 1372
Db 1357 KLEHETAVTVSEEVSKQLQTVNVPIIDGAEVSSLEGSPPCLGQEAHVCTKIQVOSSE 1416
Qy 1373 TTVTQTAABAEKV1--ETWVISEGESPEVCYGAHLLPAEKSSATGCHWTLQHAEDTVPGLG 1430
Db 1417 ASFTLTAABAEKVLGETANILETGETLEPAGAHVLEEKSSKEDNDFAAHPGEDAVPTG 1476
Qy 1431 PESQAESPIIIVTAPBSTLHPDLOGEISASQERSBEEDKPDAGPDADGKESTAISKVL 1490
Db 1477 PDCQAKSTPVIVSATTKKGLSSDLEGEKTTSLKWKSDSEVDQVACQEV--KVSVAIEDL- 1533
Qy 1491 KAEP--ILELESKNKVLNVIQTAVDQPART--ETAPETHAYDSQTVQVACRLDSREP 1547
Db 1534 --EPENGILELETSSKLVQNIQTAVDQFVRTETATEMTLSLTQTAHVIKADSDQAG 1591
Qy 1548 RCWTK-----MKDAMKHVPVQPPREDLQVLTVLEA 1577
Db 1592 QETEKEGEEPQASQADETPITSKEESESSTAVGQA 1626
```

RESULT 11

```
US-09-738-877-3
; Sequence 3, Application US/09738877
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan
; APPLICANT: Weils, Stephen J.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF ANGIOGENESIS, COMPOSITIONS, AND MET
; TITLE OF INVENTION: SCREENING FOR ANGIOGENESIS MODULATORS
; FILE REFERENCE: A-69806/DJB/JJD
; CURRENT APPLICATION NUMBER: US/09/738,877
```

```
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT/US 00/22061
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-738-877-3
```

```
Query Match 53.2%; Score 4294; DB 21; Length 1781;
Best Local Similarity 57.9%; Pred. No. 1-2e-238;
Matches 959; Conservative 187; Mismatches 401; Indels 108; Gaps 35;
```

```
Qy 1 MGAGSSTEQRSRPSQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 56
Db 1 MGAGSSTEQRSRPSQPPGSGSTPAEPSPGSGPSAEAPDTTADPAIAASDPATKLLQNG 60
Qy 57 QLSVNGVAGQDVHVQENQEQ-----EEVVDEDVQRESESDVREKOR 102
Db 61 QLSTINGVAGQDLSLQEGDLNGKQALNGQGALNGQEEVEVITVEVQDRDSDVSRDS 120
Qy 103 VEEAANSTAVEDITTKGQOETSEIIQIIPASNNVVEMVQPAESQANDVGFKKVRFVG 162
Db 121 DKEMATKSAVVDITDDQENRN-TEQIPSSSENLELTQPTESQANDIGFKKVFVFG 179
Qy 163 FKFTVKDKNEKSDTVOLLTVKKDEGEBAEASVAGAGHQBPSPVETAVGESAKSESLKQS 222
Db 180 FKFTVKDKTEKPTVQLLTVKKDEGEBA--AGAGDHQDPSL--GAGEAASKESEPKQS 234
Qy 223 TEKQEGTLKQEQSTETPLQAESDQAAEBAKQEBEKEKPTKSPESPSSPVNSSTTS 282
Db 235 TEKPEETLKQEQHAETSPPAESQAV-ECKEGBEKEKQEPKSAESPTSPVTSETGS 293
Qy 283 SFKFFFTGHGAGWKRTSPKSKEDDLETAERKEQBAEKVDEEKEKTEPASEE----- 337
Db 294 TFKKFTQGWAGWKRTSPKPKEDVEASEKKQEBEPEKVDTEEDGKAIVASEKLTASE 353
Qy 338 ----QEPADTDQARLSADYKVELPLEDQVGLDEASSEEKCAPLATEVEFDEKMEAHQ-E 392
Db 354 QAHPQBPASAEHPRLSAEYKVELPSEEQVSGSQGPSEBKPAPLATEVEFDEKIEVHQEE 413
Qy 393 VVAEVHVSTVEKTEBEGGGBAGGVVVEGTGESLPPEKLAEPQEVPOBAEAEELMKS 452
Db 414 VVAEVHVSTVEERTEQ-----KTEVEETAGSVPAEELVGMADAPQBAEPAKELVKL 465
Qy 453 REMCVSGDDHTQLTDLSPREKTLPKHPEGIVSEVEMLSOERIKVOGSLPKLFPSSSLGLK 512
Db 466 KETCVSGEDTQAGADLSPEKLVSKPGEVSEVEMLSOERMVKVQSLPKLFTSTGLK 525
Qy 513 KLSGKKQKGRGGRGDEEPGEYQIHTESPESADEQKGESSASSPEEPETTCLEKGPLE 572
Db 526 KLSGKKQKGR--GGGDEESGEHTQVPADSPDSQEQKGESSASSPEEPETTCLEKGLAE 584
Qy 573 APQGBAEBGTTSDGKKREGITPWASFQKQWTPKKRVRRPSPSPGSGEGVSTWESFKRL 632
Db 585 VQDGEAEAGATSDGKKREGVTPWASFQKQWTPKKRVRRPSPSPGSGEDKELDKVKSATLSS 644
Qy 633 TDSVTSEMQDEVKTVGEQKPEEPKRRVDTSVSWEALICVGSKKKARKAKSSDDEGPR 692
Db 645 TESTASVQEMKGSVEEPKPEPKKVDTSVSWEALICVGSKKKARRRRSSDEEGPK 704
Qy 693 TLGDSHRAEEASKQKEAGTDVAPASTQEQDQAGSSSPSPGSGEGVSTWESFKRL 752
Db 705 AMGDDHQKADGAKDXTGTDGLAGSQEHDPQGGSSSPQAGSPTEGEGVSTWESFKRL 764
Qy 753 VTPRKSKSLEEKABD---SSVEQLSTETEPSREESWVSIKKFIPGRRKKRADGKQEQ 808
Db 753 VTPRKSKSLEEKABD---SSVEQLSTETEPSREESWVSIKKFIPGRRKKRADGKQEQ 808
```

```

Db 765 VTPRKXSKLEEKSEDSIAGSVESHSTPDTPEKGEESWISIKFTIPGRKKRPRDQXQ 824
Qy 809 ATVEDSGVEINEDDPNPAVAVPLSEYNAVEREKME--AQNTLPLQLGAVVSELS 865
Db 825 APVEDAGFTGANEDSDVPAVAVPLSEYDAVEREKMEAOQAQGAQPEQKATEVSELS 884
Qy 866 KTVHTVAVAVIDGRATVSVSESPMSISATVTEPLETATGAEAMPVEVTEKIIA-E 924
Db 885 ESQVMMMAAAGDTRAAFTIIEERSPSWISASVTEPLEVEAEALLTVEEVEREVIABE 944
Qy 925 ETPVLTOTLPEKGDADMDMTSEVDFTSSEAVATETSEALRTVEEYVTEASGAETIDMVA 984
Db 945 EPPVTEPLEPKEAREAGDTVVSSEALTEPAVTAETAGLGESEGTESAABETTEMVA 1004
Qy 985 VSQITDSPTTEATPVOEVESGVLDTSEERQTAIIQAVADKYVESQVAT----- 1038
Db 1005 VSQITDSPTTEATPVOEVESGVLDTSEERQTAIIQAVADKYVESQVAT----- 1064
Qy 1039 -QTVORTSKALEKVEVEDESEVLASEKEDVMKGPVQEGAEHLAOGSETQATPES 1097
Db 1065 LQPVORA-----EAREPEEQEASGLKKEETDVILKVAQEKTEPFTQGVVGGTTPES 1118
Qy 1098 LE-VEEVTADVDH---VATQCV-----IKLOOL-MEOAVAPESSETLTPSETNGSTPLAD 1147
Db 1119 PEKAPQVTESSISELVTTCQAEITLAGVKSQEMWEOAIPDSVETPTDSETDSTPVAD 1178
Qy 1148 SPTADGTODETITDSODSKATAVAVSQVTEBEAATAQKEEPTLPNNVPAQEEHGEERG 1207
Db 1179 PPAQPTQODELVEIHEENEVAVSGTQSGTEAENAVPAQKERPPA-PSSEVFQETKEQSK 1237
Qy 1208 -RDVLEPTQOELTAAVAVPLATEVGOEGEVMLDGEKTK-----EEOVFVHSG--PMS 1259
Db 1238 MEDTLEHTDKEVSVETVSLISTEGTQ--EADQVADKTKDVPFEGLEGSIDGTIVSR 1295
Qy 1260 QCAADVUTDSEWVAGVAGCOEKESTEVOS-----LSLEGEMETDVEKEKREKTPQOVSE 1314
Db 1296 EKVTEVALKGEETEBEACKDQDALBEOHAKAPPSPEVESEMVOVHERETEMEPHVNABE 1355
Qy 1315 G-EOETAAPEHGIYQKPYLLTDMPSERKALGSLGGS-PSLPDQKAGCIEVOVQSLD 1372
Db 1356 KLEHETAVTSEVSEVSKQLQTVNVPIDGAKVSLSESGPPCLQGEAEVCTKIYQVQSE 1415
Qy 1373 TTVVTOAEAVEKVI--ETVVISETGESPEVCVGAHLLPAKKSATGHWTLQIAEDTPRG 1430
Db 1416 ASFTTLTAAEEKVLGETNMLLETGETLEPAHAHLVEKSSSEKNEDEFAHAPGEAVPFG 1475
Qy 1431 PPSQAESEPIITPAPESTLHPDLQGEISASQERSEEDKPDAGDADGKESTALEKVL 1490
Db 1476 PPOQAKSTVIVSATTKKLSLSDLEGEKTSLSKMSDEVDEQVACOEV--KVSVAIEDL- 1532
Qy 1491 KNAPE--ILEESKSNKIVLAVIQTAVDQPART-ETAPETHAVDSQTVPAQRDLSREBN 1547
Db 1533 --EPENGILELETSSKVLQNTIIQTAVDQFVTEETATMLTSELQTOAHVAKDSQDAG 1590
Qy 1548 RCWTK----MKDAMKAPVPOPREDLQVLYLEA 1577
Db 1591 OETKEGEEPPQASADEPTPITSAKERSESTAVGOA 1625

```

```

; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-13

Query Match      53.2%; Score 4294; DB 23; Length 1781;
Best Local Similarity 57.9%; Pred. No. 1,2e-238;
Matches 959; Conservative 187; Mismatches 401; Indels 108; Gaps 35;

Qy 1 MGASSSTHQRPEP-PAGSDPTSELVLSHGCPAAEAS-GAAGPADAD--DPATLPQXG 56
Db 1 MGASSSTHQRPEP-PAGSDPTSELVLSHGCPAAEAS-GAAGPADADADTADALIASPDTLLKNG 60
Qy 57 QLSNVGVAEGQDVHVEENQEQ-----EEEVDEVDVGOSEEDPVEKDR 102
Db 61 QLSITNGVAEODELSLOEGDNGQKALNGOGALNSQEBEVITYEVGQRDSEVSEKRS 120
Qy 103 VEEMANSTAVEDITTKQOEBTSEIIEOIPASENNVEMVQPAESQANDVGFKKVFRVG 162
Db 121 DKEMATSAVVDITDDQGEENRN-IEQIPSSSENLBELNQPTESQANDLIGFKVFRVG 179
Qy 163 KFTVKKDKNEKSTVOLLTYKDEGEGAEASVAGCHOEPSVTAVGESAKSELSKOS 222
Db 180 KFTVKKDKTEKPTVOLLTYKDEGEGA--AGAGDHODPSTL--GAGEAASKESEKOS 234
Qy 223 TEKQEGTLKQESSTELPLQESDQAAEBEAKDEGEKQEKPTKSPSPSPNSSETTS 282
Db 235 TEKQEBTLKQESHAELISPPAESQAV-EECKEGBEKEKPEKSKASBSPVTSSETGS 293
Qy 283 SFFKFTHGAMGRKKTFSFKSKEDLETAERKKEQBAEKVDEBEKTEBPASBE----- 337
Db 294 TFKKFTQGMAGMRKKTFSFKKEDLETAERKKEQBAEKVDEBEKTEBPASBE----- 353
Qy 338 ----QEPADTDQARLADYKVELPLBDYQGLDFASSSEKCAPLATVPEKMEANQ-E 392
Db 354 QAHQEPAPESHAEBRLSLAEYKVELPSEBQVSGSQGSEEPALALAEVPEKLEIHOEB 413
Qy 393 VVAEHAHVSTVEKTEBEOGGGEAAGVVEEGTGESLPEPKLAEPQOEPQEPAPAEIAMS 452
Db 414 VVAEHAHVSTVEKTEBEO-----KTEVEETAGSVPAEELVGMADAEPOEAPKEIYKL 465
Qy 453 REMCVSGDHTQTLTDLSPBEKTLPKHPGIVSEVEMLSQOERIRYQOSPCLKLPSSSGLK 512
Db 466 KETCVSGEDPTQADLSPDEKVLKSPREGVYSEVEMLSQOERIRYQOSPCLKLPSSTGLK 525
Qy 513 KLSGKKQKRGKGGGDEPGEYQHIHIESPESADQKGESSASSPEPEEETTCLEKPLE 572
Db 526 KLSGKKQKGRK-GGDDESGEHTQVPADSPDOEKQESSASSPEPEEETTCLEKGLAE 584
Qy 573 APQGEAEEGTTSQGEKKREGITPMASFKKVTPKKVRPRSESDKEBELEKVSATLSS 632
Db 585 VQOGEAEAGATSGEKKRREGVTPMASFKKVTPKKVRPRSESDKEBELEKVSATLSS 644
Qy 633 TDSVSEMODEVKTGVEQKPEEPKRRVDTSVSWEALICVSSKKRARKASSSSDDEGPR 692
Db 645 TESTASQEMQEMKGSVEEPKPEEPKRVDTSVSWEALICVSSKKRARRRSPDEBGP 704
Qy 693 TLGDSHRAEASDKREAGTDAVASTOEDQAOAGSSSPBAGSPSREGGVSTWESFKRL 752
Db 705 AMGCDHQADADAGDKETGDIILAGSEHPCQSSSPBAGSPSREGGVSTWESFKRL 764
Qy 753 VTPRKXSKSLEEKAEAD-----SVVEQLSTELTEPSEESWWSIKFTIPGRKKRPRDQXQ 808
Db 765 VTPRKXSKSLEEKSEDSIAGSVESHSTPDTPEKGEESWISIKFTIPGRKKRPRDQXQ 824
Qy 809 ATVEDSGVEINEDDPNPAVAVPLSEYNAVEREKME--AQNTLPLQLGAVVSELS 865
Db 825 APVEDAGFTGANEDSDVPAVAVPLSEYDAVEREKMEAOQAQGAQPEQKATEVSELS 884

```

```
Qy 866 KTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGBAMPPVSEVTEKDIITA-E 924
Db 885 ESQVHMMAAADGTRAAITIEERSPSWISASVTEPLEQVEAEALITTEVLEEVIAEE 944
Qy 925 ETPVLITOTLEGKADHDMVTSEVDFTSSEATATETSEALRTEVTEASGAETTDVMSA 984
Db 945 EPPVTEPLENREARGDTTVSEAELETPAETAETAGPLGSEGTETASAAETTEMVSA 1004
Qy 985 VSQLTDSPTDTEETATPQVESGVLDTEEBEROTQAILQAVADVKVEESQVPAT----- 1038
Db 1005 VSQLTDSPTDTEETATPQVEGGVPDIEEQERRTOEVLOQVAEKVEESQPLGTPGPEDV 1064
Qy 1039 -OTVORTGSKALEKVESEVEDSEVLASEKDKVMPKGPVQVQVQVQVQVQVQVQVQV 1097
Db 1065 LQPVQRA-----EAEPEPQAEASGLKGTDDVVLKVDQAQEAETPTQGVQVQVQV 1118
Qy 1098 LE-VPEVTADVH---VATQV-----KLQQL-MEQAAPVSESSETLTDSETNGSTPLAD 1147
Db 1119 FEKAPQVTESESELVTTQAEFLAGVKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1178
Qy 1148 SDTADGTQDQDETIDSQDSKATAAARVQVTEEEAATAQKEEPSTLPNNVPAQAEHGE 1207
Db 1179 FDPAGTTQKDEIVEIHEENEVASGTQSGGTEAEAVPAQKERPPA-PSSFVQEBTEQSK 1237
Qy 1208 -RDVLEPTQOELTAAPVPLAKTEVQGEVQVQVQVQVQVQVQVQVQVQVQVQVQV 1259
Db 1238 MEDTLEHTDEKVSVEVTSIILSKTEGTQ--EADQVADKTKDVPPFEGLEGSIDTITVSR 1295
Qy 1260 QKAADVTVYDSEVMVAGCOEKESVEVOS----LSLEEGEMETDVEKEKETREPEQVSEE 1314
Db 1296 EKVTEVALKEGTAEACKDDALELOSHAKSPSPVERMVQVQVQVQVQVQVQVQVQVQV 1355
Qy 1315 G-EQETAAPHEGYTGPVLTLDPSSERKALGSLGS-PSLPDQDKAGCIEVQVQSLD 1372
Db 1356 KLEHETAVTVSEEVSKQLQTVNVPIDGAKVSVSSGSPPCPLQGEAEVCTKIQVQVSE 1415
Qy 1373 TTVTQTAABAEKVLI--ETVVISSETGESPEVCVGAHLLPAEKSSATGCHWTLOHAEDTVPLG 1430
Db 1416 ASFTLTAAAEKKEKVLGETANILETGETLEPAGAHVLEEKSSSEKNEVFAAHPGSDAVPTG 1475
Qy 1431 PESQAESIPITVPAPESTLHPDLQGISASQERSERSEEDKPDAGPDAGKESATIEKVL 1490
Db 1476 PDCQAKSTPVLVSATTKKGLSSDLEGKTTSLKWKSEDEVQVACQEV--KVSVAISDL- 1532
Qy 1491 KAEP--ILELESKSNKIVLNVITAVDQFART-ETAPETHAYDSQTVQVACRLDSREP 1547
Db 1533 --EPENGILELETKSKSLVQNIQTAVDQFVRTETETATEMLTSELQTAHVIKADSDAG 1590
Qy 1548 RCWTK-----MKDAMKHVPQPPREDLQVLTVLEA 1577
Db 1591 QETEKEGEPEQASQADETPITSAKESESESTAVGQA 1625
```

RESULT 13

```
PCT-US01-08631-51377
; Sequence 51377, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 51377
; LENGTH: 1795
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: DOMAIN
; LOCATION: (415)...(463)
; OTHER INFORMATION: G-protein gamma subunit profile domain identified by eMATRIX,
; OTHER INFORMATION: accession number BL50058, p-value=5.557e-09, raw score of 27.23
; NAME/KEY: misc feature
; LOCATION: (1)-(1795)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-51377
```

Query Match

Best Local Similarity 52.8%; Score 4262.5; DB 1; Length 1795;
Best Local Similarity 57.5%; Pred. No. 8.2e-237;
Matches 951; Conservative 190; Mismatches 410; Indels 103; Gaps 33;

```
Qy 1 MGAGSSTEQRSPEQ-PAGSDTPSELVLSGHGPAEAS-GAAGDPADA--DPATKLPQKNG 56
Db 9 MGAGSSTEQRSPEQPSGSGTAPAEPPSGGSAEAPDTTADPAIAASDPATKLLQKNG 68
Qy 57 QLSNVNGVAEOGDVHVQENQEQ-----BEEVDEVDVQGESDVEKDR 102
Db 69 QLSTINGVABQDELSQLQEGDLNGKALNGQALNSQEEVEIVRQVDHIDSDVSDRDS 128
Qy 103 VEWMAANSTAVEDITKDGQDETSEIIQIOPASBNVNMVQPAESQANDVGFKKVFKFVG 162
Db 129 DTEWTKSAVVHDITDDQDETPIIIQIPSSSENLEELTQPTESQANDIGFKVFKFVG 198
Qy 163 FKFTVKDKNKEKSDTVQLLTVKKDEGEGAPASVAGDGHQEPSVETAVGESASKESELQS 222
Db 189 FKFTVKDKTEKPDVTQLLTVKKDEGEA---AGAGDHQDPSL--GAGEAASKESEPKQS 243
Qy 223 TEKQEGTLKQEQSSTIPIQAESDQAAEAKDEGEKQEKTKPSPESSPVNSETTS 282
Db 244 TEKPEETLKREQSHARISPPAESQAV-EECKEKEGEKQEKPSKSAESTSVTSETGS 302
Qy 283 SFKFFTHGWAGWKTSFKKSKEDDLETAEKKEQAEKVEDEEKEKTEPASEE----- 337
Db 303 TFKKFTQGWAGWKTSFRPKEDVEASEKKKEQBEKVDTEEDGKAARVASEKLTASE 362
Qy 338 ----QEPADTDQARLSADYKVELPLEQVGDLEASSEEKCAPLATEVFEDKMEAHQ-E 392
Db 363 QAHQPQPAESAHEPRLSAEYKVELPSEEQVSGSQGSEKPAPELATEVFDEKIEVQEE 422
Qy 393 VVAEVHVTVEKTEBEGQGGGAEAGVVGEGTGESI-PPEKLABPQVPPQBAEABELMKS 452
Db 423 VVAEVHVTVEERTTEQ-----KTEVEETAGSVPAEELVGMADAEPOEABEPAKLVKL 474
Qy 453 REMCVSGDHTLTDLSPREKTLPKHPEGIVSEVMLSQERIKVQGSPLKLPSSSGLK 512
Db 475 KETCVSGEDPTQGDADLSPEKVLSPKPPGVSEVMLSQERMKVQGSPLKLTSTGLK 534
Qy 513 KLSGKKQKQKGGGDEEPEGYQHIHTESPESADEQKGESSASSPEPEETTCLEKGPLE 572
Db 535 KLSGKKQKQKGR--CGGDEESGEHTQVPADSPDSQEQKGESSASSPEPEBITCLEKGLAE 593
Qy 573 APDQGAEBEGTSDGKKREGITPWASFKKMVTPKRVRRPSSDKEEELKVKVKSATLSS 632
Db 594 VQDQGEAEAGATSDGKKREGVTPWASFKKMVTPKRVRRPSSDKEDELKDKVKSATLSS 653
Qy 633 TDSVTSEMDDVKTVEGQKPEEPKRRVDTSVSWEALI CVGSSKKRARKASSSSDDGGPR 692
Db 654 TESTASEMOEEMKGSVEEPPKPKKVDTSVSWEALI CVGSSKKRARRGSSSDESGGPK 713
Qy 693 TLGGDSHRAEASAKDEAGTDAVPASTQEQDQAGSSSSPEPAGSPSEGEVSTWESFKRL 752
Db 714 AMGDDHQKADAEAGKXETGTDGILAGSQEHDPQGGSSSPQAGSPTEGEVSTWESFKRL 773
Qy 753 VTPRKSKSKLEEKABD----SSVEQLSTEIPEPSRESVWSIKKFTIPGRKKRAGDKQEQ 808
Db 774 VTPGKSKSKLEEKSDSLAGSVEHSTPDTEPKESVWSIKKFTIPGRKKRAGDKQEQ 833
Qy 809 ATVEDSGPVEINEDDNPVAVPLSEYNAVEREKME---AQGNTELPOLLGAVVSEELS 865
Db 834 APVEDAGPTGANEDDSDVPAWPLSEYDAVEREKMEAAQQAQKSAEQPEQKAATEVSKELS 893
```

```

QY 866 KTLVHTVAVAVIDGTRAVTSVEERSPSMISASVTEPLEHTAGEMPVYEVEKXIIA-E 924
Db 894 ESQVHMAAAVADGTRAAITIIERSPSMISASVTEPLEVEEAALLTEVELEKREVIABE 953
QY 925 ETPVLTQTLPECKADHDMVTSEVDFTSSEAVATATESEALRTVEVTEASGAETTDMSA 984
Db 954 EPPVTEPLPEKREARGDVTSEALTEPAVTAETAGLGAEGTEASAAETTEMVSA 1013
QY 985 VSQLTDSPTTEATPVQEVESGVLDTSEERQTAIIQAVADKYKESQVPAT----- 1038
Db 1014 VSQLTDSPTTEATPVQEVESGVLDTSEERQTAIIQAVADKYKESQVPAT----- 1073
QY 1039 -QTVQRTSKALEKVEVEDESEVLASEKEKDMKPGVQEGAGAHLAGSGSTGQATPS 1097
Db 1074 LQPVQRA-----EAERPEEQAEASGLKKEVTVLVKDAQAEKTEPFTQKVVGGTTPBS 1127
QY 1098 LE-VEEVTADVDH---VATCCV-----IKLQOL-MEQAVAPESSETLDTSETNGSTPLAD 1147
Db 1128 PEKARQVETESIESSELVTTCCQAFITAGVKSQEMWMOAIPPDSVETPTDSETDSTPVAD 1187
QY 1148 SDTADGTQODETIDSQSKATAVAVRQSVTEBEAATAQKEEPTLPNNVPAQEGHEBEPG 1207
Db 1188 FQAPGTQODEIVEIHEEBEVAASGTGCGTEAAVPAQKERPPA-PSSFVFQEEETKEQSK 1246
QY 1208 -RDVLEPTQOELTAAVPLATTEVQGEVDMIDGEKTK-----EEQEVFVHSG--PMS 1259
Db 1247 MEDTLEHTDEKVSVEVTSILSTKEGTQ--EADQVADKTKYDVPFEGKEGSLDTGTVSR 1304
QY 1260 QCAADVATDSEWVAGVACQEKESTEVOS-----LSLEGEMETDYEKRETK-DEQVSE 1313
Db 1305 EKVTEVALKGEETEBEACKDDBALBLOSHAKSPSPVEEEMVQVYERREYQXQSPITHVE 1364
QY 1314 EG-EOETAAPEHEGTYGKPVLTLDMPSSERKALSGSGS-PSLPDQXAGCIEVQVOSL 1371
Db 1365 EKLHEHTAVTVEEVSQKLOITVNPVIDAKESVSLSGSPPCQGEBAVCTKIQVQSS 1424
QY 1372 DTTVQTAAVAVKVI--ETVIVSETGESPECVGAHILPPEKSSATGCHTLOAHBTVPVL 1429
Db 1425 EASFLITAAEBEKKVIGETANILETGETLEPPGAHLVLEKSSKKNEDPAHGPGBAVPT 1484
QY 1430 GRESQAESIPITVTPAPESTLHPDLOGEISASORSEEDKPDGPDADGKESTAIKCV 1489
Db 1485 GPCQAKSTPVVVSATTKAGLSLDEGEKTSLSKMSDEVDQVACQER--KVSVAIEED 1542
QY 1490 LKAEBEILELESKSNKIYLVNIQTAVDQFART-ETAPETHAVDSQTVPACSLDSEBPNR 1548
Db 1543 LEPENGILLETFSKSLQVNIQTAVDQFVTEETATEMLTSELQTAHVIRADSDAQ 1602
QY 1549 CMTK-----MKDAKMHPPVQREDLQVTVLEA 1577
Db 1603 ETEKEGEDPQASADEPTITSAKESESTAVGQA 1636

RESULT 14
US-60-245-228-307
; Sequence 307, Application US/60245228
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
; FILE REFERENCE: C1000878
; CURRENT APPLICATION NUMBER: US/60/245,228
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 630
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 307
; LENGTH: 1779
; TYPE: PRT
; ORGANISM: HUMAN
US-60-245-228-307
Query Match 52.7%, Score 4253, DB 27, Length 1779;

```

```

Best Local Similarity 57.6%; Pred. No. 2,9e-236;
Matches 952; Conservative 184; Mismatches 406; Indels 110; Gaps 33;

QY 1 MGAGSSTBORSPEQ-PAGSDTPSELVLSGHGPAAPASGAADPADA-----DPATLQ 53
Db 9 MGAGSSTBORSPEQPPGSSSTPAEPSPGGGPMAL-----AAPDPAAETIAMDPAATLQ 65
QY 54 KNGGLSSVNGVABEGDVHVEENQEQ-----EEBVNVEDVQORESEVRE 99
Db 66 KNGGLSTINGVAEDELSTLQEGDNLNGKALNGALNSQEEBEVITVEGQORSEVSD 125
QY 100 KDRVEMAAANSTAVADITDKQGEETSEIIIEQIPASENNVMEVWPASQANDVDFKTVFK 159
Db 126 SDK--EWAATSAVVDHITDQGEETPELIRQIPSESENLELTQPTSQANDIGFKVFK 183
QY 160 FVGFKFTYKDKNEKSDTVOLLTVKQDGBGAASVGAQDHPSPVETAVGESAKSESL 219
Db 184 FVGFKFTYKDKNEKSDTVOLLTVKQDGBGA--AGAGHODPSL--GAGEAASKESEP 238
QY 220 KQTEKQEGTLKQHOSSTELPLQAESDQAAEEBAKDGEEKQEKPTKSPSPSPNSE 279
Db 239 KQTEKPEETLKROSHAEISPPAESQAV--EBCKEGEEKQEKPSKASPTSPVTS 297
QY 280 TTSEFFKFFTHGAMGMRKKTSPFKSKEDDLETAERKQEAKEKDEBEKTEBPASER-- 337
Db 298 TGSITPKFFTHGAMGMRKKTSPFKPKEDVAVASERKKEQEBEKDTEDEGAEVASEKLT 357
QY 338 -----QEPADTDQARLSADYKVELPLEDYQDLBASSEEKCAPLATEVPEKMEAH 390
Db 358 ASEQAHPEPAPESAHPEPLSLAEYKVELPSEBQVSGSQGPSEEPAPLATEVPEKIEVH 417
QY 391 Q-EVVAEYHNVTEKTEBEOGGGABEGVVEGTGSLPPEKLAEPQVQEAEPAEEL 449
Db 418 QEBVAEYHNVTEKTEBEOGGGABEGVVEGTGSLPPEKLAEPQVQEAEPAEEL 469
QY 450 MRSREMCVSGDHTQTLDSPEEKLTPKHPREGIVSEYEMLSQERIKVQSGPLAKLTFSS 509
Db 470 VKLKECTVSGSDPTQAGLSDPEKYLAKPREGIVSEYEMLSQERIKVQSGPLAKLTFST 529
QY 510 GLKLSGKKQKQKGGGGDEEPEGVQHIHTSPSADQEKSESSASPEEPEETTCLEKG 569
Db 530 GLKLSGKKQKQK--GGGDEESGHTQVPADSPQOEKQSESSASPEEPEETTCLEKG 588
QY 570 PLEAPQDEBAEBGTTSDQEKREGITTPASFKKAVTPPKRRRSPESQEBELKVSAT 629
Db 589 LAEVOQDEBAEBGTTSDQEKREGITTPASFKKAVTPPKRRRSPESQEBELKVSAT 648
QY 630 LSTSTVSEMODDEVKTYGEEQKPEEPKRRVDTSVSEALICVSSKKRARRAKASSDDEG 689
Db 649 LSTSTVSEMODDEVKTYGEEQKPEEPKRRVDTSVSEALICVSSKKRARRAKASSDDEG 708
QY 690 GPTLIGDSHRAEASKQKEAGTDAVPASTQEDQAGSSSPPEPAGSPSEEGVSTWESF 749
Db 709 GPKAMGQDHQKADKQKKTGTGDIILAGSQEHDPGQSSSPPEQAGSPTEGEVSTWESF 768
QY 750 KRLVTPRRKKSXKLEKED-----SVFQLSTELPSPSEBWSIKKTRIPRRKRRADGK 805
Db 769 KRLVTPRRKKSXKLEKEDSISAGSVHSTPDEPGKEBWSIKKTRIPRRKRRADGK 828
QY 806 QEOATVEDSGPVEINEDPNPVPVPLSEYNAVEREKE---AQNTPELPOLLGAVVSE 862
Db 829 QEOAVEADAGPTGANEDSDVPVAVPLSEYDAVEREKEAQAQKSAQPOKQATVSK 888
QY 863 ELSKTLVHTVAVAVIDGTRAVTSVEERSPSMISASVTEPLEHTAGEMPVYEVEKXII 922
Db 889 ELSEEQVHMAAAVADGTRAAITIIERSPSMISASVTEPLEHTAGEMPVYEVEKXII 948
QY 923 A-EETPVLTQTLPECKADHDMVTSEVDFTSSEAVTATITSEALRTVEVTEASGAETTD 981
Db 949 AEEBPTVTEPLPENRR--DTVVSEALTEPAVTAETAGLGAEGTEASAAETTEM 1006
QY 982 VSAVQLTDSPTTEATPVQEVESGVLDTSEERQTAIIQAVADKYKESQVPAT--- 1038

```

Db 1007 VSAVSQTLTSDPTTEATPVQVEGGVPIIEQERRTOEVLQVAEKVEESQLPGTGP 1066
Qy 1039 ----QTVORTGSKALEKVEEVEEDSEVLASEKEKDVMPKGPVQVAGAEHLAQSGSETQAT 1094
Db 1067 EDVLQPVQRA-----EAERPEEQAEASGLKKTDDVLKVDQAQEAKEPFTQGVVGGTT 1120
Qy 1095 PESLE-VPEVTADVH---VATCQV-----IKLQOL-MEQAVAPESSETLTDSTNGSTP 1144
Db 1121 PESSEKAPQVTTESLESSELVTTCQAEITLAGVKSQEMVMEQAIPDPSVETPTDSETDGTSTP 1180
Qy 1145 LADSDTADGTQODDETISQDSKATAAARQSOVTEEEAATAQKEPSTLPNNVPAQEEHGE 1204
Db 1181 VADFDPTCTQKDEIVEIHEENEVASGTQSGTEAAVPAQKERPPA-PSFVFQEEETKE 1239
Qy 1205 EPG-RDVLPTQOELTAAAVPVLAKTEVQGEVDWLDGKVK-----EEQEVFVHSGPN 1258
Db 1240 QSKMEDTLEHTDKEVSVETVSIKSTEGTQ--EADQVADKTKDVPFPEGLEGSDITGIT 1297
Qy 1259 SOKAADVTYDSEVMGVAGCQEKESTEVQS-----LSLEEGEMETDVEKEKETKPEQVSE 1313
Db 1298 VSREKVTGEGTEEAECCKDDALELQSHAKSPSPVEREMVQVEREKTEAEAPETHVNE 1357
Qy 1314 EG-EQETAAPHEGTYGKPVLTLDMPSSERKALGSLGGS-PSLPDODKAGCIEVQVQSL 1371
Db 1358 EKLEHETAVTVSEVSKQLLTAVNPVIDCAKEVSSLEGSPPCLGQEEAVCTKIQVSS 1417
Qy 1372 DTTVTQTAABAEVKVETVVISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVPLGP 1431
Db 1418 EASFTLTAABAEKVLGETANITGETLEPAGAHVLVEKSEKNEDEFAAHPGEDAVPTGP 1477
Qy 1432 ESQAESPIIIVTPAPESTLHPDLOGETISASORERSEEDKPDAGDADGKESIAIEKVLK 1491
Db 1478 DCQAKSTPVIIVSATTKKGLSSDLEGEKTTSLKWKSDDEVQVACQEV--KVSVAIEEDL- 1534
Qy 1492 AEPEILELESKNKIVLNVICTAVDQFART-ETAPETHAYDSOTQVPACRLDSREPNCRW 1550
Db 1535 -EPENGILETKSSKLQVNIITQAVDQFVTEETATEMLTSELQTAHVIKADSDAQOET 1593
Qy 1551 TK-----MKDAKMKHPVPQPREDLQVLTVLEA 1577
Db 1594 EKEGEDPQASAOQDETPTITSKESESESTAVGQA 1625

RESULT 15
US-09-758-442-562
; Sequence 562, Application US/09758442
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM031
; CURRENT APPLICATION NUMBER: US/09758,442
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 710
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 562
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (162)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (164)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

; LOCATION: (464)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-442-562

Query Match 9.7%; Score 780; DB 21; Length 732;
Best Local Similarity 38.6%; Pred. No. 5.8e-36;

Matches 221; Conservative 87; Mismatches 220; Indels 44; Gaps 17;

Qy 1039 QTVORTGSKALEKVEEVEEDSEVLASEKEDVMPKGPVQVAGAEHLAQSGSETQATPESL 1098

Db 3 QPVQRA-----EAERPEEQAEASGLKKTDDVLKVDQAQEAKEPFTQGVVGGTTPESEF 56

Qy 1099 E-VPEVTADV---DHVATCQV-----IKLQOL-MEQAVAPESSETLTDSTNGSTPLADS 1148

Db 57 EKAPQVTTESIXSSELVTTCQAEITLAGVKSQEMVMEQAIPDPSVETPTDSETDGTSTPADF 116

Qy 1149 DTADGTQODDETISQDSKATAAARQSOVTEEEAATAQKEPSTLPNNVPAQEEHGEHPG- 1207

Db 117 DAPGTTQKDEIVEIHEENEVASGTQSGTEAAVPAQKERPPA-PSXFXQOETKEQSKM 175

Qy 1208 RDVLEPTQOELTAAAVPVLAKTEVQGEVDWLDGKVK-----EEQEVFVHSG--PNSQ 1260

Db 176 EDTLEHTDKEVSVETVSIKSTEGTQ--EADQVADKTKDVPFPEGLEGSDITGITVSRE 233

Qy 1261 KAADVTYDSEVMGVAGCQEKESTEVQS-----LSLEEGEMETDVEKEKETKPEQVSEEG 1315

Db 234 KVTEVALKGEGBTEAECKDDALELQSHAKSPSPVEREMVQVEREKTEAEPTHVNEEK 293

Qy 1316 -EQTAAPEHEGTYGKPVLTLDMPSSERKALGSLGGS-PSLPDODKAGCIEVQVQSLDT 1373

Db 294 LEHETAVTVSEVSKQLLTAVNPVIDCAKEVSSLEGSPPCLGQEEAVCTKIQVQSEEA 353

Qy 1374 TVTQTAAAEVKVI--ETVVISSETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVPLGP 1431

Db 354 SFTLTAABAEKVLGETANITGETLEPAGAHVLVEKSEKNEDEFAAHPGEDAVPTGP 413

Qy 1432 ESQAESPIIIVTPAPESTLHPDLOGETISASORERSEEDKPDAGDADGKESIAIEKVLK 1491

Db 414 DCQAKSTPVIIVSATTKKGLSSDLEGEKTTSLKWKSDDEVQVACQEV--KVSXAIEEDLE 471

Qy 1492 AEPEILELESKNKIVLNVICTAVDQFART-ETAPETHAYDSOTQVPACRLDSREPNCRW 1550

Db 472 PENGILELETKSSKLQVNIITQAVDQFVTEETATEMLTSELQTAHVIKADSDAQOET 531

Qy 1551 TK-----MKDAKMKHPVPQPREDLQVLTVLEA 1577

Db 532 EKEGEDPQASAOQDETPTITSKESESESTAVGQA 563

Search completed: December 13, 2002, 00:22:46

Job time : 339 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:13:41 ; Search time 85 Seconds
(without alignments)
1253.621 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSTEQRSPQAPGSDT.....AWAQRKCLRLQLKAPVSK 1596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237906 seqs, 6676526 residues

Total number of hits satisfying chosen parameters: 237906

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4335	53.7	1844	5	US-09-724-676-61346
2	4335	53.7	1844	5	US-09-724-676A-61346
3	4294	53.2	1781	1	PCT-US02-04915-44
4	4098	50.8	1685	5	US-09-724-676-61347
5	4098	50.8	1685	5	US-09-724-676A-61347
6	429	5.3	2665	6	US-10-203-138A-11282
7	401.5	5.0	4368	5	US-09-724-676-73643
8	401.5	5.0	4368	5	US-09-724-676A-73643
9	400	5.0	2642	5	US-09-724-676-55851
10	400	5.0	2642	5	US-09-724-676A-55851
11	400	5.0	2722	5	US-09-724-676-55858
12	400	5.0	2722	5	US-09-724-676A-55858
13	400	5.0	2915	5	US-09-724-676-55849
14	400	5.0	2915	5	US-09-724-676A-55849
15	400	5.0	2915	5	US-09-724-676-55850
16	400	5.0	2915	5	US-09-724-676A-55850
17	400	5.0	2929	5	US-09-724-676-55846
18	400	5.0	2929	5	US-09-724-676A-55846
19	400	5.0	2940	5	US-09-724-676-55847
20	400	5.0	2940	5	US-09-724-676A-55847
21	400	5.0	2940	5	US-09-724-676-55848
22	400	5.0	2940	5	US-09-724-676A-55848
23	400	5.0	2995	5	US-09-724-676-55866
24	400	5.0	2995	5	US-09-724-676A-55866
25	400	5.0	2995	5	US-09-724-676-55867
26	400	5.0	2995	5	US-09-724-676A-55867

27 400 5.0 3009 5 US-09-724-676-55863 Sequence 55863, A
28 400 5.0 3009 5 US-09-724-676A-55863 Sequence 55863, A
29 400 5.0 3020 5 US-09-724-676-55864 Sequence 55864, A
30 400 5.0 3020 5 US-09-724-676-55865 Sequence 55865, A
31 400 5.0 3020 5 US-09-724-676A-55864 Sequence 55864, A
32 400 5.0 3020 5 US-09-724-676A-55865 Sequence 55865, A
33 388 4.8 1325 6 US-10-203-138A-12693 Sequence 12693, A
34 382 4.7 2014 5 US-09-724-676-49425 Sequence 49425, A
35 382 4.7 2014 5 US-09-724-676A-49425 Sequence 49425, A
36 380.5 4.7 2101 5 US-09-724-676-49404 Sequence 49404, A
37 380.5 4.7 2101 5 US-09-724-676-49405 Sequence 49405, A
38 380.5 4.7 2101 5 US-09-724-676-49419 Sequence 49419, A
39 380.5 4.7 2101 5 US-09-724-676-49420 Sequence 49420, A
40 380.5 4.7 2101 5 US-09-724-676-49421 Sequence 49421, A
41 380.5 4.7 2101 5 US-09-724-676-49422 Sequence 49422, A
42 380.5 4.7 2101 5 US-09-724-676-49423 Sequence 49423, A
43 380.5 4.7 2101 5 US-09-724-676-49424 Sequence 49424, A
44 380.5 4.7 2101 5 US-09-724-676A-49404 Sequence 49404, A
45 380.5 4.7 2101 5 US-09-724-676A-49405 Sequence 49405, A

ALIGNMENTS

RESULT 1

US-09-724-676-61346
; Sequence 61346, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 61346
; LENGTH: 1844
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61346

Query Match 53.7%; Score 4335; DB 5; Length 1844;
Best Local Similarity 58.2%; Pred. No. 2.1e-248;
Matches 962; Conservative 188; Mismatches 401; Indels 102; Gaps 32;

Qy 1 MGAGSTEQRSPQAPGSDTPESELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQNG 56
Db 62 MGAGSTEQRSPQAPGSDTPESELVLSHGHPAAEADPTTADPAIAASDPATKLLQNG 121
Qy 57 QLSSVNGVAEQGVHVGQENQEQ-----EEVDEVDYQRESEDVREKOR 102
Db 122 QLSTINGVAEQDELSQLQGLNGQKALNGQALNSQEEBEVITVEVGQRSDVDSEK 181
Qy 103 VEEMAANSTAVEDITKQGEETSEIIQIPASENNVEMVQPAESQANDVCFKKVFPVG 162
Db 182 DKEMATKSAVVHDTTDDQGEETPEIIQIPSESNBELTQPTESQANDIGFKVFPVG 241
Qy 163 FKFTVKDKNEKSDTVOLLTVKKDEGGAASVAGAGHQBPSVETAVGESAKSELSKQS 222
Db 242 FKFTVKDKNEKSDTVOLLTVKKDEGGAASVAGAGHQBPSVETAVGESAKSELSKQS 296
Qy 223 TEKQEGTLKQESSTETPLQAESDQAAEEBAKDEGEKQEKPTKSPESPSPVNSSETTS 282
Db 297 TEKPEETLKQESHAELSPPAESQAV-EECKEBSGEKQEKPSKSAESPTSPVTSKTS 355
Qy 283 SFKKFPFTHGAWMKKTSFKKSKEDDLLETAKEKQEAQKVDDEEKEKTEPASEE----- 337
Db 356 TFKKFFFTQGWAGWRKTSFRKPEDEVEASEKKKEPEKVDTEEDGKAEVASEKLTASE 415
Qy 338 ---QEPAEETDQARLADYKVELPLEDVGDLSESEKCAPLATEVDFEKKAEHQ-E 392
Db 416 QAHQPEPAESAHPRLSAEYKVELPSEBQSGSQSPSEKPAPLATEVDFEKKAEHQ-E 475

Qy	393	VVVVHVHVTYKTEKEBEOGGGGGABAGGVVAVGTESSJ.PEPKLAPEQVPOEAPAEKLKMS	452
Qy	453	REMCVSGSDHTOLTLDSPEPEKTLPRKPEGI.VSEVEVMSLOERIKYQSGPLKLKFPSSGLK	512
Db	476	VVAHVHVSIVBERTEEO-----KTEVEETIGSVPAEVLVMDAEPQAPAEVLK	527
Qy	528	KETCVSGEDDPTQADLSPDEKVLSKPEGVVSEVMTLSSQEHMKYQGSPLKLKLPSTGLK	587
Qy	513	KLSGKQKQKRGGGGGDEPQBYQHITHESPEADQKSSSSASSPEBPETTLCKEGLPE	572
Db	588	KLSSKKQKQKGR-KGGDEESEHQTQVPADSPDQOEKQSSASSPEBPETTLCKEGLAE	646
Qy	573	APQGEABEGTSSGCEKKRGITPMASFCKMVPFKRVVRPESPDEBELEKVCASATLSS	632
Db	647	VQOQGEABEGATSSGCEKKRGVTPMAFPMKMTVPKRVVRPESDKEDBLDVKYKATLSS	706
Qy	633	TDSTVSEMODEVKVGEQKPEEPKRRVDTSVSMELICVSSSKRRARAKASSSDDEGPR	692
Db	707	TESTASSEMQEMKQSVBEPKPEEPKRRVDTSVSMELICVSSSKRRARAGSSDDEGPRK	766
Qy	693	TLGGDSHRAEBASKDXEAGTDAVPASTQEOBDAQSSSPEPAGSPBEGEGVSTWESFKRL	752
Db	767	AMGEGHQADKADGAKDKETGTDGILAGQEHDPQGSSSPEQAGSPTEGCGVSTWESFKRL	826
Qy	753	VTPPKKSKSLKEEAED-----SSVEOLSTIEBSEBSWVSIKKLPGRKKRRADQKQEO	808
Db	827	VTPPKKSKSLKEEKSEDSIAGSGVGHSTPTEGKKEBSWVSIKKLPGRKKRRPDQKQEO	886
Qy	809	ATVBSQGVLENEDPNVPAVVP.LSEKNVAREKME---AQGNTELPOLLGAVVSELS	865
Db	887	APVEDAGPTGANEEDSVPAVVP.LSEIDAVERKMAQQAQASADQEPKAAITVSKLS	946
Qy	866	KTLVHTVSAVIDGTRAVTSVEERPSWISASVTEPL.EHTTAGBAMPVEVTEKDIIL-E	924
Db	947	ESQVHMAAAVAADGTRAAITIIIEBSPSWISASVTEPLEQVEAEMALLTEVLEREVIAEB	1006
Qy	925	ETPVLTOTLPRGKAHDWMTSEVDPFISEAVNTITSESLRTRIEVTVASGAETTDWMSA	984
Db	1007	EPFVTEPLPENRRARDVTVSEELTPEAVTAEFAGPLGAEEGTEAABAETTEMVSA	1066
Qy	985	VSOLTSDPTTEBEATPVOEVESGLVLTDEEEROTQALLOAVADKYVEESQVAT-----	1038
Db	1067	VSOLTSDPTTEBEATPVOEVBGVPOLEBEERITQEVLOAVAEKKEBSQLPRTGGPBDV	1126
Qy	1039	-QTVORTGSKALEKVEEVEEDSEVLASEKEDVMPKQVQEGAEHLAOGSETQCATES	1097
Db	1127	LQPVQRA-----EAERPEQAEASGLKKEITDVLKVDQAEKTEPFTGKVAVQGTDES	1180
Qy	1098	LE-VPEYTAVDH--VATQV-----IKIQOL-MEQAVABESSSETLDTSENGSTPLAD	1147
Db	1181	FEKAPVOTESJESSELVTTQOAEITLAGVKSQOEWMVQOALPPSVTEPDTSETDSTPAD	1240
Qy	1148	SDTADGQOODETIDSGSKATAVVRQSVTEEBAATAQKEBPTLPNNVPAQEBEGEPG	1207
Db	1241	FDABCTTQKDIYVIEHEENEVAOSTGSGTGAEAVPAQKPRPA-PSSVVFQEEPKESK	1299
Qy	1208	-RDVLEPTQOELTAAAPVLAKEITVQEGEVDMLDDEKTK-----EEQEVFVHSG--DNS	1259
Db	1300	MEDTLEHDKESVSEVYJSLSKTEGTO--EADQYADKTKVDVPRFEGLEGSLDTGITYSR	1357
Qy	1260	QKADVTVYDSFWMGVACCQEKESTEVOS-----LSLEGBMETVDEKKEKTEKPEQVSEB	1314
Db	1358	EKVTEVALKGEGTEAEACKDKDALTELQSHAKSPSPVBEREMVQVVEREKTEAPEPHVNEB	1417
Qy	1315	G-BOETAPAPHEGTYGKRVULTLMPSSERKALGSIQGS-PSLPRDDKAGCTLEVQOSLD	1372
Db	1418	KLEHETATVYSEEVSKOLLQTVANVPIIDAKKEVSSLEGSPPLQGEAEVCTKLOVQSSB	1477
Qy	1373	TTVTOIAEAVKVL--ETVVISGESSEPCVGAHLLPAEKSSATGCHWTLQJADPTVPLG	1430
Db	1478	ASFLUTAAEBEKVLGETANILLETGETLEPAGHVLVEKSSKXEDPRAHNGEAVPTG	1537
Qy	1431	PEQAESPIIIVTAPASTLHPDQEGISASQERSEEBDKPDAGDAGKESTAIKVL	1490

```

Db      1538 PCCQAKSRPFIYVATSATTCKGLSDSDEGEKTTSLKMKMSDEVDQVACQEV--KTSVAIEEDL 1595
Qy      1491 KAEPILEILESKNSKIYLVNIQTAVDQFART-ETAPETHAYDSQTOVPACRLDSREPNRC 1549
Db      1596 EPEENILELETETKSKLTVONIIQTAVDQFVRETEATATEMLTSELQTAHVIKADSDQACQE 1655
Qy      1550 WTK-----MKQAKMHPVQPREDDQVLTVLEA 1577
Db      1656 TEKEGEBPQASQADETPITSACESESESTAVGQA 1688

RESULT 2
US-09-724-61346
/ Sequence 61346, Application US/09724676A
/ GENERAL INFORMATION:
/ APPLICANT: Compugen LTD
/ TITLE OF INVENTION: Variants of alternative splicing
/ FILE REFERENCE: 129181.4 Compugen
/ CURRENT APPLICATION NUMBER: US/09/724.676A
/ NUMBER OF SEQ ID NOS: 97222
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 61346
/ LENGTH: 1844
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-724-676A-61346

Query Match      53.7% Score 4335; DB 5; Length 1844;
Beet Local Similarity 58.2%; Pred. No.2.1e-246;
Matches 962; Conservative 188; Mismatches 401; Indels 102; Gaps 32

Qy      1 MGAGSSTGGRSPDEQ-PAGSDTPSELVLVSGHGPAEAS-GAADDPADAA--DPATKLPQKNG 56
Db      62 MGAGSSTGGRSPDEQPEGSSTPAEPPEPGSGGSAEAAPTTADPALAASDPATKLLQKNG 121
Qy      57 QLSVNGVACGQDVAVQENQEQ-----EEVVDVEDVQGRSEDDVREKOR 102
Db      122 QLSVNGVACGQDELVLQEGDLNGQALNGGALNGGALNSOEHEEVIYVTEVGGQGRSEDDVREKOR 181
Qy      103 VEEMANSAVAVDITRKDQGEETSEIIEQIPASENVVEEYVQAEQANDVGRKVPKPVG 162
Db      182 DKEMTKSAVVAHDITDDGQGEETPEIIEQIPSESNLIEELTQTESQANDIGFKKVPKPVG 241
Qy      163 FKFYTKKQKNEKSDPVOLTVTKKDGEGEAASVAGADHQPSEVETAVGSESASESELKOS 222
Db      242 FKFYTKKQKTEKPTVQLLTVKKDGEGEA--AGAGDHQDPSL--GAGAAAKESSEPKOS 296
Qy      223 TEKQGTLKQBOSSTEIPIQAESDQAEBEAKDQGBEKOKEKPTYSPESSPVNSSETTS 282
Db      297 TEKPEETLKREQSHAEISPPAESQAV-EECKEGBEKEKEKPEKSAESPPTSSETGS 355
Qy      283 SKKKFPTQMAWRKKTSPFKSKEDDLETAERKKEQAEKVDEEKEKTEPASEE----- 337
Db      356 TEKKFPTQMAWRKKTSPRKPEDEVESEKKKQEPKVTDEGKAQVASEKLTASE 415
Qy      338 ----DEPAEDTQARLSADYEKVELPLEDOVDLPAASSEKCAPLATEVFDKMEAHQ-E 392
Db      416 QAHPEPRAESAEPRLSAEYKVELPSEQVSGSGSPSEKAPLATEVFDKIEIVHQHE 475
Qy      393 VVAEYVSTVEKTEEBEGGGGAEAGVNVETGESLPPEKLAEPQEVPOEAPAEELMKMS 452
Db      476 VVAEYVSTVEERTBEQ-----KTEVEETAGSVPABELVEMADPEAPAEKELVTL 527
Qy      453 REMCISGCDHTLTLDSPEKTLPKHPBGIVSEVEMLSQGEIKIYQGSPLKLLPSSGKL 512
Db      528 KETCVSGEDPTGADALSPKEKVLSKRPEGVASEVEMLSQGEIMKQVGSPLKLLPSTGKL 587
Qy      513 KLSGKKQKKRGGGDEEGEYQHTHTESPEADBEQKGSASSPPEPPTTCLKGPLE 572
Db      588 KLSGKKQKKGR-GGGDEEGEHTQVPADSPDQBEQKGSASSPPEPPTTCLKGLAE 646

```

```
QY 573 APOGEAEEGTTSGEKKBEIGITWASFKKMWTPKKVRRPSESDBKEELEKVKSATLSS 632
Db 647 VQDGEAEEGATSGEKKREGVITWASFKKMWTPKKVRRPSESDBKEDLDKVKSATLSS 706
QY 633 TDSVSEMQDVKVTGEBEQPEEKRRVDTSVSWEALICVGSKKRKARKASSDDEGGPR 692
Db 707 TESTASEMQEMKGSVEEPKPEEKRRVDTSVSWEALICVGSKKRRARRGSSDEEGPK 766
QY 693 TLGDSHRAEASKDKBAGTDVAVPASTQEQDAQGSSPPEAGSPSGEGVSTWESPKRL 752
Db 767 AMGDHQADEAGKDKETGTDIGLAGSQEHDPGCGSSSPQAGSPTEGEGVSTWESPKRL 826
QY 753 VTPKKSLSLEAKED----SSVEQLSTELPESRESWSIKKFIQRRKKRAGDKQEQ 808
Db 827 VTPRKSLSLEAKESDISAGSGVEHSTPDTEPKGSEWSIKKFIQRRKKRPGDKQEQ 886
QY 809 ATVEDSGPEINEDDPNVAVPLSEYNAREKME--ACGNTELPOLLGAVVSVBELS 865
Db 887 APVEDAGTGANEDSDVPVAVPLSEYDAVEREKMEAAQQAQSAEQPEQKAATEVSKELS 946
QY 866 KTLVHTSVVAVIDTAVTSVEERSPSWISASVTEPLEHTAGTAEAMPPVEEVTEKDIIA-E 924
Db 947 ESQVHMAAAVADGTRAATIEERSPSWISASVTEPLEQVEAEAAALLTEEVLEREVIABE 1006
QY 925 ETPVLQTQLPEGKADHDMVTSEVDFTSSEAVTATETSEALRTEBTEVTSAGABETDMVSA 984
Db 1007 EPPVTPELPENREARGDVTVVSEALTEPVAATAETAGPLGAEGTEASAAEETEMVSA 1066
QY 985 VSQLTDSPTDTEETATPQVESGVLDTEEBERQTOALQAVADVKVEESOVPAE----- 1038
Db 1067 VSQLTDSPTDTEETATPQVEGGVPDIEEQRRTOEVLQAVAEKVKEESQLPTGGPEDV 1126
QY 1039 -QTVORTGSKALEKVEEVEDSEVLASEKEKDVMPKGPVQDAGAEHLAQGSETQCATPES 1097
Db 1127 LQPVORA-----EAERPEQAESGLKETTDLVVKDAQEAKEPTPTQGVKVCQTTPES 1180
QY 1098 LE-VPEVTAQVDH---VATQV-----IKLQOL-MEQAVAPESSETLTDSETNGSTPLAD 1147
Db 1181 FEKAPQVTEGIESSELVTTTQCAETLAGVKSQEMVMEQAIIPDSVETPTDSTDGSTRVAD 1240
QY 1148 SDTAGTQDQETIISQDSKATAAARQSVQTEEAATAQKEPSTLNNVPAQEBHGEPG 1207
Db 1241 FDPCTTQKQELVIEIHENEVASGTQGGTEAEAVPAQKERPPA-PSFVFQEBTKQSK 1299
QY 1208 -RDVLEPTQOEBLTAAPVLAKTEVGOGEVDWLDGKVK-----EQEVFVHSG--PNS 1259
Db 1300 MEDTLEHTDKSVSVETVSIILSKTEGTQ--EADQYADEKTKDVFPFEGLEGSIDGTIVSR 1357
QY 1260 QKAADVTYDSEVMGACQOEKSTEVQS-----LSLEGENMETDVEKEKRETKPEQVSEE 1314
Db 1358 EKVTEVALKGEGTEBEACKDDALELQSHAKSPSPVEREMVQVVEREKTEAEPTHVNEE 1417
QY 1315 G-EQETAAPHEGTYGKPLVLTDMPPSSERKALGLSGS-PSLPDQKAGCIEVQVQSLD 1372
Db 1418 KLEHETAVTVEEVSQKLLQTVNVPIDGAKESVSLSGSPPCGQGEAEVCKLQVQSSE 1477
QY 1373 TTVTQTAABAEKVJ--ETVVISSETGESPECVGAHLLPAEKSSATCGHWTLOHABDTVPLG 1430
Db 1478 ASFTLTAABAEKVLEGTANILETGETLEPAGAHVLEKSSERKNEDFAHPGEDAVPTG 1537
QY 1431 PESQAESPIIIVTAPESTLHPDLOGEISASQERSSEEDKPDAGPADGKESITAEKVL 1490
Db 1538 PDCQAKSTPVIIVSATTKGLSSDLEGETTSLKWSDEVDRQVACQEV--KVSVAIBEDL 1595
QY 1491 KAEPEILELESKSNKIVLNVITQAVDOPART-ETAPETHAVDTSOTVPACRLDSREPNC 1549
Db 1596 EPENGILELETSSKLVONIIQTAVDOPVTRTEETATENMTSELQTAHVIKADSQDAQOE 1655
QY 1550 WTK-----MKDAMKHPVQPPQREDLQVLTVLEA 1577
Db 1656 TEKGEFPOASAQDETPTSKESESTAVGQA 1688
```

RESULT 3

```
PCT-US02-04915-44
; Sequence 44, Application PC/TUS0204915
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200PC
; CURRENT APPLICATION NUMBER: PCT/US02/04915
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/285,475
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-04915-44
```

```
Query Match 53.2%; Score 4294; DB 1; Length 1781;
Best Local Similarity 57.9%; Pred. No. 5.4e-246;
Matches 959; Conservative 187; Mismatches 401; Indels 108; Gaps 35;
```

```
QY 1 MGAGSSTEQRSPBQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 56
```

```
Db 1 MGAGSSTEQRSPBQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 60
```

```
QY 57 QLSVNGVAEQGVHVOEENQEQ-----EEVVDVEDVQSESDVREKDR 102
```

```
Db 61 QLSTINGVAEQDELSQEGDLNGKQALNGQALNSQEEVEVITVEVGQDSEDVSRDS 120
```

```
QY 103 VEWMAANSTAVEITKQGOSETSEIIQIPASNNVEMVQPAESQANDVGFKKVFPVG 162
```

```
Db 121 DKEMATKSAVVHDITDDQGENRN-IEQIPSSSNLEELTQPTESQANDIGFKKVPKVG 179
```

```
QY 163 FKFTVKDKNEKSDTQVLLTVKKDEGEBAASVAGAGDHQPSVETAVGESAKSESELKQS 222
```

```
Db 180 FKFTVKDKTEKPTVQLLTVKKDEGEBA--AGAGDHQDPSL--GAGEAASKESEPKQS 234
```

```
QY 223 TEKQEGTLKQEQSSTEIPLQASDAQAAEBAKQEGEKQKPTKSPSPSPVNSSTTS 282
```

```
Db 235 TEKPEETLKREQSHAETSPPAESQAV-ERCKEGBEKEQKSPKSAESPTSPVTSETGS 293
```

```
QY 283 SFKKFFTHGAGWRKTKTSFKKSKEDDLTAETAEKEQBAEKVDEEKEKTEPASEE----- 337
```

```
Db 294 TFKKFFTHGAGWRKTKTSFKKSKEDDLTAETAEKEQBAEKVDEEKEKTEPASEE----- 353
```

```
QY 338 ---QBEPAEDTDQARLSADYKVELPLEDQVGLDLEASSEEKCAPLATEVDFDKMEAHQ-E 392
```

```
Db 354 QAHPQPAESAHAEPRLSABEYKVELPSEEQVSGSQSPSEKPAFLATEVDFDEKIEVHQEE 413
```

```
QY 393 VVAEVHVSTVTEKTEBQGGGEAEGGVVVEGTGESLPPEKLABEPQEVPOBAEPAELMKS 452
```

```
Db 414 VVAEVHVSTVTEETEEQ-----KTEVEETAGSVAEELVGMADBPQAEPAKELVKL 465
```

```
QY 453 REMCVSGDHTQTLTDLSPBEKTLPKHPEGIVSEVEMLSQERIKVQGSPLKLPSSSGLK 512
```

```
Db 466 KETCVSGEDPTQAGDLSPDEKVLSPKPPGVVSEVEMLSQERMKVQGSPLKLPSTGLK 525
```

QY 513 KLSGKQKQKGGGDEPGEYOHITHSPESADOKBSSASSPEPEBETTCLKGPLE 572
 Db 526 KLSGKQKQKGR -GGGDESGEHTQVPADSPQOEQKSSASSPEPEBETCLEKGLAE 584
 QY 573 APQGEAEEGTTSOBEKREGITTPWASFKKWTPKKRVRRPESKPEBELEKVKATLSS 632
 Db 585 VQOQGEAEEGATSDSEKREGEVTPWASFKKWTPKKRVRRPESKPEBELEKVKATLSS 644
 QY 633 TDSITVSEKQDEKTVGEEQKPEEPKRRVDTSVSEWALICVSSSKKRAKASSDDEGGPR 692
 Db 645 TESTSEKQDEKKSVEEPKPEPKRKYDTSVSWALLICVSSSKKRAKARRSSDDEGGPR 704
 QY 693 TLGDSHRAEASKDKKXAGTDAVPASTQODQACSSSPPEAPGSGSEGGSTWMSFKRL 752
 Db 705 AMGCHQKADKXGDKKXGTDGILAGSQEHDPQSSSPPEQAGSPTEBEGVSTWMSFKRL 764
 QY 753 VTPRRKSKSLEEKAD-----SSVQLSTELPESREESVSTKKFIPGRRKKRADKQKQ 808
 Db 765 VTPRRKSKSLEEKSEDSIAGSGVHSTPDTPEKESVSTIKKFI PGRRKKRPDQKQ 824
 QY 809 ATVEDSGVEINEDDPNPAVPLSEYNAVEREKME--AQNTLPLQLLGAVYVSELS 865
 Db 825 APVEDAGTGANEDSDVPVAVPLSEYDAVEREKMEAQKAGAQPEQKATEVSKELS 884
 QY 866 KTLVHTSAVAIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEKDTIA-E 924
 Db 885 ESQVHMAAAVADGTRAAITIIERSPSWISASVTEPLEVEAALITIEVLEVERVIAEE 944
 QY 925 EFPVLTOTLPEKKAHDMMVTFSEVFTSAVATATSEALRTETEYTRASGAEETDMVSA 984
 Db 945 EFPVTEPLPENREARGDTIVSEALTEPRAVTAETAGPLGSEEGTSAABETTEMVSA 1004
 QY 985 VSQLTDSPTTEAPVQVESGVLDETEEERQTOAILQAVADKVKESQVAT----- 1038
 Db 1005 VSQLTDSPTTEAPVQVEGEGVLDIEQERQTOEVLQAVAKKESQGLGTGGEDV 1064
 QY 1039 -QTVORTSKALEKVEVEEDSEVLASEKQVMPKGPVQEAHLAGHLAGSTGQATPES 1097
 Db 1065 LQPVQRA-----EABREPEQASGLKKTEDVVLKVDQAEKTEPFGKAVGQTPES 1118
 QY 1098 LE-VEEVNADVH---VATCOV-----IKLQOL-MEQAAVAPESSETLJDSNNGSTPLAD 1147
 Db 1119 FEKAQVVESTISSSELVTTCQAEITLAGVKSQEMVMEQALPPDSVETPTDSEDTGSPVAD 1178
 QY 1148 SPTADQOODETIDSDSKATAAVROSQVTEEAATAQKEPSTLPPNNYPAQEBEGBG 1207
 Db 1179 FPARGTQKDEIVEIHEHEVNASGQSGTEAFAVPAQKERBPA-PSFVFEETKEGSK 1237
 QY 1208 -RDVLEPTQOELTAAAVPVLATVEGQGEVMDLGEKVK-----EBQEVFVHSG--PNS 1259
 Db 1238 MEDTLEHTDEKESVETVSTILSTKEGTQ--EADQVADKTKDVPFEGLEGSIDTGTYSR 1295
 QY 1260 QKAADVTVDSSEWAGVAGCOEKSESTVOS-----LSLEGEEMETDVEKEREKTRPEOVSE 1314
 Db 1296 EKVTETALKEBETSEBECKKDALBELQSHAKSPSPFVEREMVVOVEREKTEAPHTVANE 1355
 QY 1315 G-EQETAAPEHEGTGKPLVTLDMPSSEKKAIGSLGSG-PSLPPQDKAGCIEVQVQSLD 1372
 Db 1356 KLEHETAVTVSEVSKQLQTVNVPITIDAKVSSLSGSPPCLOGEAEVCTKIQVQSSSE 1415
 QY 1431 PPSQASBPITITVPAPESTLHPDLOGEISASQRESEEDKADGADAGDKSTAEKVL 1490
 Db 1476 PDCQAKSTPVIIVSATTKGLSSDLBEKKTSLKWSDEVDQVACQEV--KVSVALBDL- 1532
 QY 1491 KAEPE--IIELEKSKKIVLANTVQTAVDQFART-ETAPETHAYDSOTQVPAKRLDRBN 1547
 Db 1533 --EPENGILLETSKSLVONITIIQTAVDQFVITEETATMLTSELQTAHVIAKDSQDAG 1590

QY 1548 RCMWK-----MKDAKMKHPVQPREDLQVLTVLEA 1577
 Db 1591 QETKEGEBEPQASQADDETPITSAXEBSSESTAVGQA 1625
 RESULT 4
 US-09-724-676-61347
 ; Sequence 61347, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 61347
 ; LENGTH: 1685
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-61347
 Query Match 50.8%; Score 4098; DB 5; Length 1685;
 Best Local Similarity 58.4%; Pred. No.2,1e-234;
 Matches 904; Conservative 177; Mismatches 382; Indels 84; Gaps 28;
 QY 89 VQGESESDVREKDVEMANSTAVEDITYKQGEETSEIIIFQIPASENNVEMVQPAESQ 148
 Db 9 VQGRDSEVSESDSDKEMATKSAVAVHIDTDCGQETFEIIEQIPSSSINLELTQPTESQ 68
 QY 149 ANDGFKKVFKEVGFKFTVKKDKNEKSDTVOLLTVKKDEGGEAASAGAGDHQSPVETA 208
 Db 69 ANDGFKKVFKEVGFKFTVKKDKTEKPTVOLLTVKKDEGGA--AGAGDHQPSL--G 123
 QY 209 VGESAKSESELKQSTKQEGTLKQOOSTEITPLQAESDQAAEEBAKDEGEKQKEPTKS 268
 Db 124 AGEASKSESEPKQSTKQETKEETLKREQSHALISPPAESGQAN-EECKKEGEEKQKEPEKS 182
 QY 269 PESPSSPVNSFTTSPFKKFTTHAGWRKKTSPFKSKBEDDLETAKKRKEQAEKVEDEEK 328
 Db 183 AESPTSPVTSSTGTFKFFFTQAGWNRKKTSPFRKPEDEVEASEKKKEQPEKVTREED 242
 QY 329 EKTPEASE-----QEPADTDQARLSADYKVELPLEDQVGLDEASSEKQPIA 379
 Db 243 GKAEVASKLTASQOAPQEPASAEPRLSAEYKVELPEEBOVSGSQGSEKPEAPLA 302
 QY 380 TEVEDEKMEARQ-EVVAEVHVSIVTEKTEEEBQGGGEABGVVAVGTGSLLPPEKLAEPQ 438
 Db 303 TEVEDEKTEVHQAEEVVAEVHVSIVTEBRTBQ-----KTEVETAGSVPAEELVEMDA 354
 QY 439 VPQEAPEABELMKSRKEMCVSGGHTLTULDSPEKTLPKAPREGIVSEYMLSSQERIKVQ 498
 Db 355 EPOEAPEAKELVKKETCTVSGDEPTQAGADLSPDEKVLKSPREGVASEVEMLSQERMVQ 414
 QY 499 GSPKTKLFPSSGKLKLSGKQKGRGGGDEPGEYOHITHSPESADOKBSSASSPE 558
 Db 415 GSPKTKLFTSTGLKLSGKQKGR-GGGDESGEHTQVPADSPQOEQKSSASSPE 473
 QY 559 EPEETTCLEKPLLEAPQGEAEGTTSDEKREKREGITPWASFKKWTPKKRVRRPESBDK 618
 Db 474 EPEETTCLEKGLAEVQDQGEABEGATSDGEKKRKGCVTPWASFKKWTPKKRVRRPESBDK 533
 QY 619 EEELEKVKATLSTDSITVSEKQDEKTVGEEQKPEEPKRRVDTSVSEWALICVSSSKR 678
 Db 534 EDELKVKATLSTSTESTASQEMQKSVVEPKPEPKRKYDTSVSEWALICVSSSKR 593
 QY 679 ARKASSDDEGGPRTLGDSHRAEASKDKKXAGTDAVPASTQODQACSSSPPEAPGSP 738
 Db 594 ARGSSDDEGGPRKAMGDHQAEDKXGTDGILAGSQEHDPQSSSPPEAGSP 653
 QY 739 EGEVSTWMSFKRLVTPPKKSKSLEEKAD-----SSVQLSTELPESREESVSTIKKFI 794
 Db 654 EGEVSTWMSFKRLVTPPKKSKSLEEKSEDSIAGSGVHSTPDTPEKESVSTIKKFI 713

QY	1084	LAOGSTGATGESLE--VEP2TADVH--VAFCOV-----IKQOL-MEQAUPESSET	1133
Db	1008	FTQCKVVGQTTBESFEKAPVTESIESSELVTTQATLACVSGQEMVBEAIPFDSJET	1067
QY	1134	LTDSENGSTPLADSDTADQODETIDSDSKATPAVROSQVTEEBATAPQKEPSTLP	1193
Db	1068	PTDSEGTSTPVADFPAPGPTTQKDEITVEIHENEVEASGTOSGGTEAEVAPQOKERPPA-P	1126
QY	1154	NNVPQOEHEGEBG--RDYLEPTQOELTAAPVPLAKTEVGEGBVMDJGKXV-----E	1247
Db	1127	SSFPVPEETKEQSKMEDTLEHTDEVSVEVTSJLSKTEGT--PADQYADKTDVPPFE	1164
QY	1248	BOEVPFHSQ--PNSOKAADVTYDSEVMGVAGCOEKESTEVOS-----LSLEEGEMETDVE	1300
Db	1185	GLEGSIDTGITVAREKTEVALKGEGBAECKDDALTELQSHAKSPSPYEREMVQVE	1244
QY	1301	KEKREKTPPEOVSEB--EQETAAPEHECTYKRPVLTLMPSSEKRAKALSGGS--PSLPDQ	1358
Db	1245	REKTEPEPHVNEKELEHETAAVTVSEVSKOLQTVANVPIIDAKEVSSLEGSPPCLGQ	1304
QY	1359	DRAGCIEVQVOSADTVTQTAEAVERKI--ETVVISETGESPEVCAGHILPAEKSSATG	1416
Db	1305	EEAVCTCKIQVOSSEASPTLTAABEKKVLGETANIILEGTLEPAGAHVLEEKSEKNE	1364
QY	1417	HMTLQHAEDTVPLGPESQASPIVTVPAPESTLHPLOGEISASQRRSEEPKPDAGP	1476
Db	1365	DFPAHDGEAVPVPQOAASTPVIATTKGJSSDLGEGKTTSLKWKSSDEVBEQVACQ	1424
QY	1477	DADGKESTAIKYLKAPELLELESKSNKIIVANIQTAVDQFART--ETAPETHAVDSQTQ	1535
Db	1425	EV--KVSVAIEEDLPENGILLETKSSKLYONITQTAVDQFVTEETATAMLSLTQTO	1482
QY	1536	VPACRLDSREPNRCWTY----MKDAMKHPVPQPREDLQVLTULBEA	1577
Db	1483	AHVYKDSODAQOETEKEGEPQASADETPTITSAKESSESTAVGQA	1529

```

RESULT 6
US-10-203-138A-11282
/ Sequence 11282, Application US/10203138A
/ GENERAL INFORMATION:
/ APPLICANT: Molecular Dynamics, Inc.
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN BT 474
/ FILE REFERENCE: PB 0004 WO 8
/ CURRENT APPLICATION NUMBER: US/10/203,138A
/ CURRENT FILING DATE: 2002-08-02
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 04 February 2000 (04.02.00)
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 26 May 2000 (26.05.00)
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 03 August 2000 (03.08.00)
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 03 October 2000 (03.10.00)
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 27 September 2000 (27.09.00)
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 21 September 2000 (21.09.00)
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 30 June 2000 (30.06.00)
/ NUMBER OF SEQ ID NOS: 15438
/ SOFTWARE: Molecular Dynamics Sequence Listing Engine
/ SEQ ID NO 11282
/ LENGTH: 2665
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:

```

```

: OTHER INFORMATION: MAP TO AL034555.2
: FEATURE:
: OTHER INFORMATION: EXPRESSED IN B1474, SIGNAL = 12
: FEATURE:
: OTHER INFORMATION: EST_HUMAN HIT: A0117052.1, EVALUE 0.00e+00
: FEATURE:
: OTHER INFORMATION: SWISSPROT HIT: P08640, EVALUE 3.00e-10
US-10-203-138A-11282

Query Match 5.3%; Score 429; DB 6; Length 2665;
Best Local Similarity 19.6%; Pred. No. 1.3e-16;
Matches 372; Conservative 270; Mismatches 755; Indels 504; Gaps 83

```

QY	9	QRPSEOPAGSD-----	-TPSELTVSGHCPAAEASG-----	-AAGPDADPAT-	49
Db	301	QVOSTEPKASDLSKLESVMKYPKPEKGLSHVEVVEKGGRLKARKHLKPEOPADGVAVD			360
QY	50	--KLPOKNGOLSSVNGVAEQGVHVOEENOGEOEBEYVDEDVGORESEDAVEKRDVEMA			107
Db	361	LEKLEAKRRFADSNLKAEKQREPKSSPEMEDARVL-----	-SKQODVSSREVI----		411
QY	108	ANSTAVEDITKQOEBTSELIIOQIPASENNVEMVQAPASQANDVGPKNVFKFGFKFTV			167
Db	412	-----LIREBEARKKPRKELIKRES-----	-KX1-----		435
QY	168	KQDKNEKSDPTVOLITVYKQDEGEAGAEASVAGCHQEPSEVETAVGESASKESELKOSTEKOE			227
Db	436	-----KLRLNLTNVASPHOCQELASISVSGSSRPSSDLOARLIGELAG-----	-ESVENOE		483
QY	228	GTLKQOBSSTFIP-----	-LOAESQAAEE- -AKDEGEKQEKPEPKSPESPSSPVNS		278
Db	484	-----VOSKPRIPSKPOLQOLVTLDOQGERDVAKNYCSLRDETPEKKSQOEKSHSVNT			538
QY	279	ET-----	-TSFFKPFTHGAMGMRKTPSFKSK- -EDDLETAKR-----		315
Db	539	EKKIDIDHTQSYRKQMEDSRKQOMEIAXSKFPKSPKKDUDVEYRSLVHVEVKRP			598
QY	316	-----KEQAEKVD-----	-EEKEKTEPASEOEPAEDTDQARLSADYEKVEL		358
Db	599	QDVTDSPPKKKRMHDVPDICTKERERNYRSSROISEDSERTGSPSVRHGSFHD--			655
QY	359	PLEQVODLEASS-----	-EEKCAP-----	-LATVWDEKMEHNGVAVEVHVS	401
Db	656	--EDPISPRLLTSKSGPKYDEKVLTVYSNITVREBSLKNPDPSSRRQOMADMKKIKLSV			713
QY	402	VEKTEEEOGGGGGAEE--	-GGVVEGTGEBSLPEPKI-----	-AEQOEVPQAEPAEE--	448
Db	714	LNSDELDNRDMSQMKQADARFDVSPFNSIIXKDSLKXGSYNDLBEQGEVPPSDSDDEGHS			773
QY	449	-----	-LMSREMCVSGDHTQTLDSPE-----	-EKLTPKHPBGIVSE	485
Db	774	HSPPASALYESRSLFLLRRREDKLERBERDELSLSERNKFFSPALDXTIIPDTKALLER			833
QY	486	VEMUS-----	-----OGRIVQOSP-----	-LKKLPSSGLKTLGS	516
Db	834	AKSLSSSRERENWSFLMDSRFANFRNNKQKEKVDASPRIPSYWYKKKKIKITDSEGRKD			893
QY	517	KKQKRGKGGGDEEPGEYOHITHTESPESADEQKGGSSASSPEPEPETTCLKGPLEAPQD			576
Db	894	KKEDHK-----	-EEEOEROLFAPRFLHSSIIFQDSKRLQHLERKEDBDIFSIRIGYKQI		948
QY	577	GEABEGTTSDEKKRBEITPMAS- FKXV-----	-TPKKVRPRSESDKREELKVKXATLS		631
Db	949	SEGANSTT-----	-DSIOEPVVLFIHSRPFELTLRMQOKEKEKQKQKEVKEKQEDTENHPKTPES		1005
QY	632	STDSIVSEMQ-----	-DEKVTGVGEQOKEEPPRRVDTSSVMEALICVSSGSKRRARXAS		684
Db	1006	APENKSDSLKTPPEVGPSTVTVTLTSARSALEKTKTGDKTVEAPLV-----	-TEEKTVETATV		1062
QY	685	SDD-----	-EGGPRTLGSDSHRAEE-----	-ASKDKRAGTDAVPASTQEDQOAOQSSSEP	733
Db	1063	SEAKPKPASEBPAPAV-----	-EQLQOVLDLPGADPDKELA--	-NMPDAVEE-----	GSSGOP 1111

QY 734 ---AGSPSEGEVSTWE-----SPKRLVTPRKKSLSLEKAE--DSSVE----- 773
Db 1112 PYLDAKPTPGASQASQNSVDPEDSTQPLSKPAQSEENEPKAPKPDADAEPDAN 1171
QY 774 ---QLSTEIEPSRESWVSIKFTIPGRKKRADQGEQATVEDSGPVEINEDDPNPAVV 830
Db 1172 QKAEAAPESQPPASED-LEVDPPVAADKPKNKRKSKTPTVO-AAAVSIVE-----KPV 1224
QY 831 PLSEYNAVEREKMAQAGNT--ELPQLGAVVVSBELSKT-----LVHTVSVAVIDGT 880
Db 1225 RKSE--RIDREKLKRSNPRGEAQKJLELMEAEKIRTRASKNSAADLHPEPSLPLSRT 1282
QY 881 RA--VTSV-----EERSPSWISASVTEPLEHTAGEAMPVEEVTEKDIIAE--ETPVL 929
Db 1283 RRRNVRSVYATMGHNRSP-----VKPEVQ-----PRVTRKRLERLEQEAANV 1327
QY 930 TQT-----LPEGKADHDMTSEVDFTSSEAVTATETSEALRTEEVTEASGAE-----ETTD 980
Db 1328 PTTPRGRPPKTRRRADEEENEKEPAETLKPPGWRSPRSOKTAAGGGPOGKKGKNEP 1387
QY 981 MVSAVSOLTSPTTEBATPVQEVESGVLDTPEEEROTQAILQAVADKVKESQVPAQT 1040
Db 1388 KVDAT-----RPEATTEVGPIGVKESMEPKAAEEBAGSQKDRDKDAGTDKNPPTAP 1442
QY 1041 VORTGSKALEKVEVEEDS-----EVLASEKEK--VMPKGPVQAGAEHL 1084
Db 1443 VE-----VWEKKPAPEKNKSKRGRNSRLAVDKSASLKNVDAVSPRGAQAAGER-- 1495
QY 1085 AQGETG--QATPESLEVPETADVHVATQVVIKLOLMQEAQVAPESSETLTDTSETNGS 1142
Db 1496 -----ESGVVAVSPKSPQKEDGLS-----SQLKSDPDVDPKEPEKEDVSASGP 1541
QY 1143 TPLADSDTADGTQOEDTIDS-----QDSKATAAARQS-----QVTEEBAA 1183
Db 1542 SPEA--TQAKQMELEQAVEHTAKLAESAASAAKYADAPGLAPEDRDKPAHQASSETELAA 1600
QY 1184 AQK---PEPSTLPNNVQAQEEH-CEE-----PGRDVLEPTQOEL-TAAAVPVLAKEV 1231
Db 1601 AIGSIINDIGEENFAPPYPGESQTDLPQAPAGAALOPSEGMETDEAVSGILETEA 1660
QY 1232 QGEQEVLDGCKEVEQEVFVHGSPNSQKAADVTYDSEVNGVAGCOEKESTEVQSLSLE 1291
Db 1661 ATSSRRPPVNAAPDS-----AGPTDKARGNSSETSHSVPEAKSGEVEVTLVRKD 1712
QY 1292 EGETMDVKEKRETKPEQVSEGEQETAAPHEGTYGKPVLTLDMPSSSRGKALGSLGG 1351
Db 1713 KGRQKTRSRKRNTNKKWA--PVESHVPESNQAQ-----ESPAANEGLTV-----Q 1759
QY 1352 SPSPDQDKAGCIEVOVQSLDITVQTA-----EAVEKVIETVVISSETGESPECVGAH 1404
Db 1760 HPEAPQEEK-----QSEKPHSTPQSCSTDLKIPSTENSQBSISVEERTPTTKASVPPD 1813
QY 1405 L-----LPAEKSSATGHTWLQHAEDTVPLGPESQAESIPIITVPA----- 1445
Db 1814 LPPPPQAPVDEEQARFRVHSIESDPVTPPSD-PSIPIPTLPSVTAALKSPVVASGGI 1872
QY 1446 PESTLHPDLOQIESASQRESSEEDKPDAGDADGKE-----STAIEKVLKAEPIELES 1501
Db 1873 PHQSPPTKVTEWITROEPPRAQSTPSPALPDTKASDVDTSSSTLRKIL-MDPKYVVSATS 1931
QY 1502 KSNKIVLNVITQAVDQFARTETAPETHAYDSQTVQACRLDSREPNCRTWKMADKMKHP 1561
Db 1932 VTSVTVTAIAEPV-----SAAPCLH-----EAPPPVDSKKP-----LEEKTA 1972
QY 1562 VPQPREDLQVLTLEANAQPKCLPRQLK-----APVS 1595
Db 1973 VTN-NSEIQASEVLVA-ADKEKAPVIAPIAKTISVISRMPVS 2011

RESULT 7

US-09-724-676-73643

; Sequence 73643, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73643
; LENGTH: 4368
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676-73643

Query Match 5.0%; Score 401.5; DB 5; Length 4368;

Best Local Similarity 19.9%; Pred. No. 1.2e-14;

Matches 317; Conservative 272; Mismatches 662; Indels 339; Gaps 63;

QY 8 EORSPPQAGSD---TPSELVLSHGHPAAEASGAAGDPADADPATKLPK-----NGQ 57
Db 1374 EERLAEQQAERERLAEVEAALEKORQLAAHAQAQAARE-AKELQORMQOEVVVRE 1432
QY 58 LSSVNGVAEOGDVHVOENQOGEVEVVDVQRESE-----DVREKDRVVEEMAANS 110
Db 1433 EAAVD--AQOKRSIQEELQOLROSSEAEIQAKARQAEEAERSRLRIEIEIRVRLQLEA 1490
QY 111 TAVEDITKQOQETSBI-IEQIPASNNVVEVMQPAESQANDVGFVKVFGVGFKFTVKK 169
Db 1491 TERQGGAGELQALRARAEAEAKQQAEEARLRQVQDESQRKQAEVELASRVKA 1550
QY 170 D---KNEKSDTVOLLTVKKDEGEAE-----ASVGAGDHQEPSVET-----AVGESAKES 217
Db 1551 ETEAAREKQALQALBELQLQAEAEERLRQAEVBARQVVALETAQRAAEAELOSRA 1610
QY 218 ELQSTQSTQKQGTLLKQOQ--SSTEIPLQAE---SQAAEEBAKDEGEKQEKETKSPSPS 273
Db 1611 SPKETAQLERSQEBHVAQUREEARRAQQAEEAAREEARELEERWQKANEALR 1670
QY 274 SPVNSSETTSFKFFPTHGWAQWRKKTSPFKSKEDDLETAKRKEQAEKVDDEBEKTEP 333
Db 1671 LRLQAEVAQOKSL-----AQAAEAKQEEAREARRRRKGAEEQA-VRORELAQE- 1720
QY 334 ASEQEPADTDQARLSADYKVELPLEQVQGLEAS--SEKCAPL-----ATEVFDE- 385
Db 1721 LEKQRLAEGTAQORLAEEQELIRLRAETEQEQOQRLLEELARLQREAAAAATQKRQEL 1780
QY 386 -----KWEAHQEVV-----AEVHVTVEKTEEGGCGEAGEGVVVEGTGE-----SL 428
Db 1781 EALAKVRAMEVILLASKARAEESRSTSEKSKQRL-----EAGAFRELAEEAARLAL 1836
QY 429 PPE-----KLAEPQEVPOEAPAEELMKSRMCVSGGDHTQLTDLSPKPTLPHPEGIV 483
Db 1837 ABEAKRQQLAEBAARQRAE-AERVLAEKLAAI--GEATRL-----KT-----EAEI 1881
QY 484 SEVEMLSQBERIKVQSGPLKPLSSSGLKLSKKQKQKRGGGDEEPGEYQIHHTESPE 543
Db 1882 ALKEKAENER-----LRLAEDEAFORRLLEEQAAQHKADIEERLAQ---LRKASDS 1931
QY 544 SADEQKQ--ESSASSPEEPEITCLKGPLEAPQDCEAE-----EGTSPGKKEKREG 593
Db 1932 ELERQKGLVEDTLURQRQVEEETILAKSFEKAAAKAELELELGRIRSNABDTLSRKEQ 1991
QY 594 ITWASFKKMVTTPKRVRRPSPESDK-----EEELKVKSATLSSTSDTSVSEMQDEBVK 646
Db 1992 AELEAARQQLAAEEERRREAREERQKSLAAEEAARQKAALEEVERLKANV-EARR 2050
QY 647 VGEQKPEPK-----RRVDTSVSWEALICVSGSKRKARFASSSDEGEPRTLGG 696
Db 2051 LRERAQESARQLQAEAAQKRLQAEKAFAPV--QQKEQLQOTLQEQSVLQDLRG 2108
QY 697 DSHRAEASKDE-----AGTDVAPASTQ-EQDOAGSSSPPEAGSPSEGEVSTWESFK 750
Db 2109 EAEARRAAEEAEARVQAREAAQARRQVEAEARLKQSAEQAAQAAQAAAA----- 2162

```

QY 751 RLVTFRKSKSLKEKADSSVEQLSTEIEPSRESWVSIKKFIPGRKKRKGADQKQCAT 810
Db 2163 -----EKLKKEAEQEAARRAQAEQALROKQADAMEKHKKFAEQTLROKQAEQELTT 2217
QY 811 VEDSGPVEINEDDPNVPVAVPLSEYNAVERE-----KME-----AQONTLPOLLGAVVYS- 861
Db 2218 LR-----LQLEETD-----HQNKLDEBELQRLKAEATEAARQSRQVSEBELFSVRQOM 2264
QY 862 EELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPV-EEVTEKD 920
Db 2265 EELSK-----LKARIEAENRALILRDKDNTORFLOEBAEKKKQVAAEBAARLS 2311
QY 921 IIAETPVLTO-----TLPEGKAHNDWTSSEVDFTESEATATETSEALRTTEVTASGAE 976
Db 2312 VAAQGAARLRQLAEEDLQAQRLAKMKLKKQAVQEAIRLKAEBELLQOQKELAOEQAR 2371
QY 977 EETDVAWSVQSQTSDPDTTEEAATPVOEVESGVLDTEEEERQATLQAVADAVK----- 1030
Db 2372 RLQF-----DQEQMAQOLAEFTQGFORTLEAERQRLMSAEBELKLRVAMEM 2419
QY 1031 EESQVPATQTVORTGSKALEKVEEVEEDSEVLASEKEKQVMEKGPVQEAHLLAQSSET 1090
Db 2420 SAAQARAEBDAQRFKKQAEIGEKHLRTE--LATQEKVLTVQTLFIOQSDHDAE---- 2473
QY 1091 GQATEESLEVEPVTAVDH--VATQVITLQOL-MEQAVAPESSETLTDSENGSTPLAD 1147
Db 2474 -----RLREAIABLEREKELQOEALTLQKSEMOTVQOEBLLOETQALQOSFSE 2525
QY 1148 SPTADGTODDEFITDSQDSKATAAARQSOVTEEBEAAATAKEBPSTLPNNVPAEENGHEEPG 1207
Db 2526 KQSL--LQERITLQEKALQ-----EQLFQDEVAQAQO-----LAEF--QOQO 2564
QY 1208 RDVLEPTQOELTAAAVPVLAKTEVGOGEVDWLDGEKVEKEOEVPVHSGPNSQKADAVTY 1267
Db 2565 QOQMEGERQLVAASMEARROHBAEG-----VARKOE----- 2598
QY 1268 DSEWNGVACQEKESKESTEVSLSIEBEMETDVKEKERTKPE-QVSEBEGQETAAPEHES 1326
Db 2599 -----ELQOLEQORQOEBELLAEENQRLREQLQLLEEQHRAALAHSEEV 2642
QY 1327 TYGKPLVLTLDMS-----SERKALGSLGSPSLPDQDCACTEV-QVOSL---- 1371
Db 2643 TASQVAATITLNGRHALDGPAAEAEPHSPDGLRKVSAQALQEGILISABELRLAQS 2702
QY 1372 DTTVTQTAEAVE-----KVLETVVISETGE-----SPECVGAHLIPAEK 1410
Db 2703 HTTVBELARREVDVRYLQGRSSIAGLLLKATNEKLSVVALQROLLSP--GTALLILEA 2759
QY 1411 SSATG-----GHWTLOHAEDTVPLGPE 1432
Db 2760 QAASGFLDPRVNRRLTVEAVEKEGVGPE 2789

RESULT 8
US-09-724-676A-73643
; Sequence 73643, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 73643
; LENGTH: 4368
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-73643

Query Match 5 0%; Score 401.5; DB 5; Length 4368;
Best Local Similarity 19.9%; Pred. No. 1.2e-14;
Matches 317; Conservative 272; Mismatches 662; Indels 339; Gaps 63;

```

```

QY 8 EORSPEQAGSD-----TPSELVLSGHGPAABASGAGDPADAPATKLPOK-----NGQ 57
Db 1374 EERLAEQRAEERERLAEVEALEKQRLABAHQAQAQABRE-AKELQOMQOEVEVVRRE 1432
QY 58 LSSVNGVAQGDVAVQENQEGQEEVVDVQGRESE-----DYREKDRVEAMANS 110
Db 1433 EAAYD--AQOQKRSIQBELQOLROSSSEAFQAKARQAEABERSRLRIEBELRVVRLQLEA 1490
QY 111 TAVEDITKQGEETSET-IEQIPASENNVEBMOVPAESQANDVGFKKVFFKVFGEKFTLVKK 169
Db 1491 TERQGAEGELQALRAAEAEAEQQAQOEAEERLRRQVODESORQQAQAEVLASRVKA 1550
QY 170 D--KNEKSDTVQOLLTVKKQEGAE-----ASYGADHDQEPSVET-----AVGSASKE 217
Db 1551 ETEAAREKQRLQALBELRLQAEERLRLOAEVERRAROVVALETQORSSEALQSRDA 1610
QY 218 ELKOSTEKQEGTLKQOQ-SSTEIPLQAE--SDQAEEEAQDEGEKQEKPTSPSPS 273
Db 1611 SPAEKTQQLERSLOEBHVAVQAQLEAEARRAQQAABEABAREAELEERWQLKANFALR 1670
QY 274 SPVNSSETTSFKKPFTHGMAGWRKTSFKSKEDDLETAERKQEAQKQDEEKEKTER 333
Db 1671 LRLQAEVVAQOKSL-----AQAEAEKQEBAREARRRGAERQAVQRELAQOE- 1720
QY 334 ASEBQEPEDTDQARLSADYKVELPLEDQVGDLEAS-SEKCAPL-----ATEVFDE- 385
Db 1721 LEKQRLAEGTAQORLAEOQLIRLRAETEGEQEOQRLLEBELRLRLREAAATATQKQEL 1780
QY 386 -----KMEAHQEVV-----AEVHSTVEKTEEEQGGGAEAGVVEGTGE-----SL 428
Db 1781 EAEALKVAABEVLASAKABEESRSTSEKQRL-----EAEAGRPRELAEABARLAL 1836
QY 429 PPE-----KLAEPQEPQEAAPAEELMKSRMVCVSGDHTQTLSPBEKTLPHNPEGIV 483
Db 1837 ABEAKRQRLAEEDAARQAE--AEVLAELKLAII--GEATRL-----KT-----EAEI 1881
QY 484 SEVENLSSQERITKYQSGSLKTLFSSSGIKLSGKKQKQKGGGDEBERGEVQHITHTEPE 543
Db 1882 ALKEKAEANER-----LRLAEDAEAFORRLLEQAAQAHKDAIEERLQO--LRKADS 1931
QY 544 SADEQKQ--ESSASSPEEPETTCLEKPLTAPQDGEAE-----EGTTSDEKKREG 593
Db 1932 ELERKGLVEDTLQRQRYEEIILAKASPEKAAAGAELELELGRIRSMNEDTLRSQO 1991
QY 594 ITWASFKKMVTPKKRVRRPESDK-----EELEKVKATTSSTDSTVSEWODEVKT 646
Db 1992 AELFAARQRLAEERERRRREAEERVQKSLAEAEBAARQKALAEVEERLKANY-EEARR 2050
QY 647 VGEQKPEEPK-----RVDTSVSWEALICVSSSKKRAKASSDDEGGPRTLG 696
Db 2051 LREAREQSSARQOLAQEAQKRLQAEKHAFAV--QOKEQLQCTTQOQOSVIDQRLG 2108
QY 697 DSHRAEASDKQE-----AGTDVAPASTQ-EQDAQSGSSPEPASPESGEGVGTWESFK 750
Db 2109 EAEARARRAAEBEERARVQAEBAQAQRQVEAEELKQSAEQQAQAAQAAAA----- 2162
QY 751 RLVTFRKSKSLKEKADSSVEQLSTEIEPSRESWVSIKKFIPGRKKRKGADQKQCAT 810
Db 2163 -----EKLKKEAEQEAARRAQAEQALROKQADAMEKHKKFAEQTLROKQAEQELTT 2217
QY 811 VEDSGPVEINEDDPNVPVAVPLSEYNAVERE-----KME-----AQONTLPOLLGAVVYS- 861
Db 2218 LR-----LQLEETD-----HQNKLDEBELQRLKAEATEAARQSRQVSEBELFSVRQOM 2264
QY 862 EELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPV-EEVTEKD 920
Db 2265 EELSK-----LKARIEAENRALILRDKDNTORFLOEBAEKKKQVAAEBAARLS 2311
QY 921 IIAETPVLTO-----TLPEGKAHNDWTSSEVDFTESEATATETSEALRTTEVTASGAE 976
Db 2312 VAAQGAARLRQLAEEDLQAQRLAKMKLKKQAVQEAIRLKAEBELLQOQKELAOEQAR 2371

```

```
QY 977 ETTWMSAVSQTDSPTTTEATPVQESGVLDTTEEROTQALLOAVADKVK----- 1030
Db 2372 RLQZ-----DKEQMAQALAEETOQFRTLAEARQRLMEGAERLKLRAVM 2419
QY 1031 EESQVPAQTQVGTGSKALEKVEEVEDSEVLASEKEDVMPKGPVQVQAGAEHLAQGSET 1090
Db 2420 SRAQAREEDAQPRKQAEIEGKLRHTE--LATQEKVTLVQTLEIQROQSDHAE----- 2473
QY 1091 GOATPESLEVPVTDVNH--VATCQVILKQOL-MEQAVAPESSETITDSETNGSTPLAD 1147
Db 2474 -----RLREATAELEREKEKLQBAKLQQLKSEEMQTVQOBLLQETOALQOCSFLSE 2525
QY 1148 SDTAGTQOQETIDSQDSKATAAARQSVQTEEEAATAQKEEPSLTPNNVPAQBEHGEPPG 1207
Db 2526 KDSL--LORERFIBQEKAL-----BOLFODEVAKAQO-----LREE--QORQ 2564
QY 1208 RDVLEPTQOELTAAAPVFLAKTEVQGEVDMDLGEKVKBEQEVFVHSGPNSQKAADVTY 1267
Db 2565 QQQMEQERQRLVASMEERARRQHEAEG-----VRRKQE----- 2598
QY 1268 DSEVMGAGQOEKESTEVQSLSLEEGEMETDVEKEKRETPQ-VQSBEGQETAAPHEGG 1326
Db 2599 -----BLQLEQOQRRQOQELLABENQRLREQLLQLEEQHRAALAHSEEV 2642
QY 1327 TYGKPVLTLDMP-----SERGKALGSLGSGSPSLPDQDKAGCIEV-QVQSL--- 1371
Db 2643 TASVAATKTLPNGRDALDGPAAEPEHSPDGLRRKVSQAQRUGAIGILSABELQRLAQG 2702
QY 1372 DTTVTQTAEAVE-----KVIEVVISSETGE-----SPECVGAHLPAEK 1410
Db 2703 HTTVDELARREDVRHVLQGRSSIAGLLLKATNEKLSVYALQORLLSP---GTALILLEA 2759
QY 1411 SSATG-----CHWTLQHAEDTVPLGPE 1432
Db 2760 QAASGFLDLPVRNRLTVNEAVKEGVGPE 2789
```

RESULT 9

```
US-09-724-676-55851
; Sequence 55851, Application US/09724676
; GENERAL INFORMATION:
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; NUMBER OF FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55851
; LENGTH: 2642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-55851
```

```
Query Match 5.0%; Score 400; DB 5; Length 2642;
Best Local Similarity 19.1%; Pred. No. 6.8e-15;
Matches 306; Conservative 304; Mismatches 663; Indels 330; Gaps 58;

QY 43 ADADPATKLPQKNGQLSSVNGVABQGVHVQEEHQEQEVEEVDVQGESEDDVREKDR 102
Db 738 AEARQRRLDYES-QTAHDNLLTFQ--IHSLSIETAKSKDKVI---EVLQNELDDVQ---- 787
QY 103 VEEMAANSTAVEDITDKGQETSIL---EQIPASENNVEEMVQPARSQANDVGFKKVKF 159
Db 788 -LQFSEQSTLIRLSQSLQNKSESVLEGAERVRIHSSKVEELSQAQSKLEI----- 839
QY 160 FVGFKFTVKDKNEKSTVQLLVTKDDEGAASVAGDHQEPBSVTAEGESAKSESEL 219
Db 840 -----TKMD-----QLLEKRDVETLQOTIEBKQQVTEISFSTKRMVQLNEE 884
QY 220 KOSTKEQGLTKQDSSTETPLQAESDQAAEEFAKDEGEKQEKPTKSPESPSSPVNSE 279
Db 885 KFSLGVEIKTLKE-----QNLNLSRABEAKKEQVEEDNEVSSGLKQNYDEMSPAQISKE 939
```

```
QY 280 TTSSFKKPFTTHWAGWRKKTTSFKKSK-----EDLETAERKRKQOEARKVDEEKEKT 331
Db 940 E-----LOHEFDLLKKEQERKRLQALINRKKELLQVRSLBELANLKDSEKKEIP 992
QY 332 EPASEQEPAPEDTDQARLS-----ADYEKVELPLEDOVGDLA-----SSBEKCAPLA 379
Db 993 LSTERGEVEEDKENKEYSEKCVTSKQETIYLKQITISEKEVELQHIRKDLSEK---LA 1049
QY 380 TEVFDKMEAHQSVAAEVHVSTVEKTEEEQGGGEAGGVVVVEGTGESLPPEKLAEPQEV 439
Db 1050 AE-----EQFQALVKQMNQTLQDKTNQ-----IDLLAQAEISENQAI 1085
QY 440 PQBAEPAEELMKREMCVSGDHTQLTD--LSPBEKTLPKHPIGIVSVEMLSQERI 495
Db 1086 IQK-----LITSNTDASGDGVALVKETVVISPPCTGSSEHMKPELEEKILALEKEKE 1138
QY 496 KVQGSPLKPLFSSSGLKLGKKGKRGGGDEEPECEYQIHITESPESADEQKGES--- 552
Db 1139 QLOKQLQEAULTSRKAILKKAQEKERHLR-----EELQKQKDDYNRLQEQDFDEQSKENENI 1193
QY 553 -----SASSPEEPETTTCKLEKPLAPQDGEABEGT----- 584
Db 1194 GDQLRQLOIQVRESIDGKLPSTDOQESCSSTPGLBEPFLPKATEQHHTTPVLESNLCPDWP 1253
QY 585 --SDGKKKREGITPMAASFKKQWTPKKRVRPSPESDKEELEKVK-SATLSSTSTSTVSEMQ 641
Db 1254 SHSEDASALQGGTSVAQIKRAQL-----KEIEAEKVELKELKVSSTTSELTKKS 1300
QY 642 DEVKTVGEBQKPEPRKRVDTSVSWEALICVGSKKRKARKASSDDGGGRTLGGSQSHRA 701
Db 1301 EEFVQLQEQINKQGLIESLKTVSHEAEVHAESLQKL-----ESSQLQIAGLEHLR 1352
QY 702 EEASKDKEAGTDVAPASTQEQDQAQGS---SSPEPAGSPSEGEVSTWESFKELVTPRKK 758
Db 1353 ELQPKLDEL---QKLISKKEEDVSYLSGQSEKALTKIQTETIIQEEDLIKALHT----- 1405
QY 759 SKSKLEKABDSVQOLSTEI-----EFSREESVWIKKFIPIGRRKKRADQKQEQATV 811
Db 1406 -QLEMQAKEHDERIKQLOVELCEMKQKPEIGES-----RAKQOIQRKLQAALI 1454
QY 812 EDGCPVEINEDDPNVAVPLSYNAVER-----EKMEAQCNTELPLLQAGVAVSEE 863
Db 1455 SRK---EALKENKSLQFELSAR-GTIERLTKSLADVESQVSAQNKEKDTVLGRLALLQE 1510
QY 864 LSKTLVHTVSAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEKDIIA 923
Db 1511 ERDKLITEDMRSLLENQSLSSSC---SLKLALEGLTDEKELVKEI-----ESLKSKSIA 1563
QY 924 EETPVLVTQTLPEGKDAHD-----MVTSEVDFTSEAVTATETSEALRTEE----- 968
Db 1564 EST-----EWQEKHKELOKEVEILLQSYENVSNEAERIQHVVAVRQEKQELYGKLR 1615
QY 969 VTASGAABETDMVSAVSQTLDSPDTTEATPVQEVESGVLDTTEEROTQALLOAVADK 1028
Db 1616 STEANKKETEKQLQEAQEQEEMKEMKRMKF--AKSQKQKILELEENEDRLRAEVHPAGDT 1673
QY 1029 VKESQVPATQTVORTGSKALEKVEVEEDSEVLASE-----KEKDVMPKGPVQAGAEH 1083
Db 1674 AKZ-----CMETLLSNASMKLELVRKMEYETLSKKFQSLMSEKQSLSE-EVQD--LKH 1725
QY 1084 LAQGSQTGAQTPESLEVPVETADVHDVATQCV---IKLQOLMEQAVAPESSETLTDSETN 1140
Db 1726 QIEGNVSKQANLEATEKHNDQNTVTEEGTQSIPGETEEQDLSMSRTPCTSESVPSSAKS- 1784
QY 1141 GSTPLADSDTAGTQOD---ETIDSDQSKATAAVROSQVTEEEAATAQKEEPSLTPNNVP 1197
Db 1785 -ANPAVSKDFSSHDEINNYLQITDQLKRIAGLEEBEKQNKESQTLNEK-NTLLSQIS 1842
QY 1198 AQEEHGEPRGVLEPTQOELT-----AAAVPVLAKTEVQGEVDMDLGEKVKEE 1248
Db 1843 TKD--GE-----LKMQLQEBVTQNNLLNQIOBELSRVTKLKTABEEDKDLLERLMNQL 1894
```


Db 788 -LQFSEOSTLIRSLQSQLOQNKSEVLEGAERVRHISKVEELSQALESKELEI----- 839
Qy 160 FVGPKFTVKDKNKSPTVOLLTVKKDEGEGAEASVAGAGDHQEPSVETAVGESAKSESEL 219
Db 840 -----TKMD-----QLLLEKRDVETLQOTIEEKDQVTEISFSMTKMWQLNEE 884
Qy 220 KQSTKQEGTLKQESSTETPLQAESDQAAEEBAKDGEBKQKEPTKSPESPSPVNSE 279
Db 885 KFSGLGVEIKTLKE-----QLNLLSRABEAKEQVEEDNEVSSGLKQNYDEMSAPAGQISKE 939
Qy 280 TTSSFKKFFTHGAWGRKKTSPFKSK-----EDDLETAERKKEQAEKVDDEEKEKT 331
Db 940 E-----LOHEFDLLKKEQERKRLQOALINRKELLQVRSLREELANLKDESKKEIP 992
Qy 332 EPASEEPEADTDQARLS-----ADYKVELPLEDQVGDLEA-----SSEKCAPLA 379
Db 993 LSETERGEVEEDKENKEYSEKVTSKQETIYLYKQITISEKEVELQHIRKDLEK---LA 1049
Qy 380 TEVPDEKMEAHQEVVAEVHVSIVTEKTEBEOGGGGEAGGVVVEGTGSLPEKLAEPQEV 439
Db 1050 AE-----EFOALVKQMNQTLQDQTNQ-----IDLLQAEISENQAI 1085
Qy 440 PQAEPAEELMKREMVCVSGDHTQLTD---LSPPEKTLPKHPEGIVSEVEMLSQERI 495
Db 1086 IQK-----LITSNTDASDGSVALVKEITVVISPPCTGSEHWKPELEKILALEKEKE 1138
Qy 496 KVGSPULKLFSSSGLKJGKKQKRGGGGDEPGEYQIHHTSPESADEQKGES--- 552
Db 1139 QLQKKLQEALTSRAILKKAQERHLR-----BELKQKDDYNRLQEQDQESKENENI 1193
Qy 553 -----SASSPEEPETTCLKGPLEAPQDGEABEGTT----- 584
Db 1194 GDQLRLQIQVRESIDKLPSTQOESCSSTPGLEELPFRATEQHHTQPVLESNLCPDWP 1253
Qy 585 --SDGEKKREGITPWASFKKWTKVRPSPESDKDEELEKYK-SATLSTDSVTSEM 641
Db 1254 SHSDEAALOGGTSVAQIKQAQL-----KEIEAKVELELKVSTTSBELTKS 1300
Qy 642 DEYKTVGEOKPEPKPRKRVDTSVSWEALICVSSKKKARKASSDDGPRTLGGDSHRA 701
Db 1301 EEFVQLQEQINKQGLEITESLTKVSHAEVHAESLQOKL-----ESSQLQIAGLEHLR 1352
Qy 702 EASKEKEAGTDVAPASTQDQQAQS---SSPEPAGSPSEGEVSTWESFKRLVTPRKX 758
Db 1353 ELQPKLDEL---QKLSKKEEDSVLSQOLSEKAALTKIQTIEIEDLILKALHT--- 1405
Qy 759 SKSKLEKAEDESSVEQLSTEI-----EPSREESVSIKKFIPGRKKRADGQEOATV 811
Db 1406 -QLEMQAKEHDERIKQLQVLECMKQKPEEIGES-----RAKQIQIRKLQALI 1454
Qy 812 EDGPFVEINDDPNPVPVPLSEYNAVER-----EKMEAQNTLPPOLLGAVVYSEE 863
Db 1455 SRK---FALKENKSLQELSRLAR-GTIERLTKSLADVESQVSAQNKEDTVLGRLLALQ 1510
Qy 864 LSKTLVHTVSVAVJGTRAVTSVEERSPFSISASVTEPLEHTAGEAMPPEVTEKDIIA 923
Db 1511 ERDKLITEMDRSLLENQSLSSCSB--SLKLALEGLTEDKEKLVKEI-----ESLKSXKIA 1563
Qy 924 EETPVLVTQTLPEGKAHDD-----WVTSEVDFTSNAVATATETSEALRTEE----- 968
Db 1564 EST-----EWQEKHKLQKEYEILLQSYENVVSNABERIQHVVEAVRQEKQBYGLKL 1615
Qy 969 VTRASGAETTTDMVASVQLTDSPTTTEATPPQVESSGLVDTEEBEERQOAILQAVADK 1028
Db 1616 STEANKKETEKLOQAEQAEQEMEKKRKF--AKSKQKILELEENDRURAEVHPAGDT 1673
Qy 1029 VKESQVPATQTVORTGSKALEKVEEVEEDSEVLASE-----KEKVMPKGPVQEGAEH 1083
Db 1674 AKE-----CMETLSSNASMKELERVKMEYETILSKFQSLMSEKDSLSE-EVQD--LKH 1725
Qy 1084 LAQGETGQATPESLEVPVETADVHVATCOV-----IKLQOLMEQAVAPESSETITDSETN 1140
Db 1726 QIEGNVSKQANLEATEKHDNQTNTVBEGTQSI PGTEBQESLSMSRPTCSVSPSAKS- 1784

Qy 1141 GSTPLADSDTADCTQOD---ETIDSQDSKATAAVROSQVTEEBAAATAQKEPSTLPNNVP 1197
Db 1785 -ANPAVSKDPSSHDEINNYLQIDQLKERIAGLEEEKQKNKFSQTLNENEK-NTLLSQIS 1842
Qy 1198 AQSEHGEESPORDVLEPTQOEBLT-----AAAVPVLAKTEVQGECEVDWLDGEKVEE 1248
Db 1843 TKD--GE-----LKMQLQEBVTQWLLNQIQBELGRVTKLKTABEEKODDLBERLMNQL 1894
Qy 1249 QEVFVHSGPNSSQKAADVT-----YDSVMGVAGQCEKESTEVSQSLSEEGEMETDVE 1300
Db 1895 AEL---NGSICNVCDQVTDQAIKNELLESEMKNLKKCVSELEBKQQLVKEKTKVSEIR 1951
Qy 1301 KEKRETKPEQVSEBEGQETAAPHEGTYGPKVLTLDMPSSERGGKALSGSGSPSLPDQDK 1360
Db 1952 KEYLEKIQGAQKEPQGNK-----SHAKELQELLKEKQEVKQL-----QKDCIRYQEK 1998
Qy 1361 AGCIEVQVQSLDITVTCATAEAVEKVIETVVISSETGESPECVGAH-----LLPAEKSSAT 1414
Db 1999 ISALERTVKALEFVQTESQKDLITKEN--LAQAVEHRKKAQAEALASFVLLDDTQSEAA 2056
Qy 1415 GGHWTLQHAEDTVPGLPESQA--ESIPITVPAPESTLHPDQCEISASOKERSEEDDKP 1472
Db 2057 -----RVLADNLKLKKELOSNKESVKSQMKQDE-----DLERRLEQAEKHLKEKNM 2105
Qy 1473 DAGPDADGKESTAIEKVL-----KAPEPILELESKSNKIV 1507
Db 2106 QEKLDALRREKVHLEETIGETIGVTLNKKDKKEVQOLQENLDSTV 2148

RESULT 14
US-09-724-676-55850
; Sequence 55850, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 55850
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-55850

Query Match 5.0%; Score 400; DB 5; Length 2915;
Best Local Similarity 19.1%; Pred. No. 7.9e-15;
Matches 306; Conservative 304; Mismatches 663; Indels 330; Gaps 58;

Qy 43 ADADPATKLPQKNGOLSSVNGVAEQGDVHVQEBEQGEVEEVDVQGESEEDVREKDR 102
Db 738 AEAEORRLDYES-QTAHDNLLTQ--IHSLSIETAKSKVKI---EVLQNELDDVQ--- 787
Qy 103 VEEMAANSTAVEDITKDGQEBTSEII---EQIPASENNVEEMVQPAESQANDVGFKKVFK 159
Db 788 -LQFSEOSTLIRSLQSQLOQNKSEVLEGAERVRHISKVEELSQALESKELEI----- 839
Qy 160 FVGPKFTVKDKNKSPTVOLLTVKKDEGEGAEASVAGAGDHQEPSVETAVGESAKSESEL 219
Db 840 -----TKMD-----QLLLEKRDVETLQOTIEEKDQVTEISFSMTKMWQLNEE 884
Qy 220 KQSTKQEGTLKQESSTETPLQAESDQAAEEBAKDGEBKQKEPTKSPESPSPVNSE 279
Db 885 KFSGLGVEIKTLKE-----QLNLLSRABEAKEQVEEDNEVSSGLKQNYDEMSAPAGQISKE 939
Qy 280 TTSSFKKFFTHGAWGRKKTSPFKSK-----EDDLETAERKKEQAEKVDDEEKEKT 331
Db 940 E-----LOHEFDLLKKEQERKRLQOALINRKELLQVRSLREELANLKDESKKEIP 992
Qy 332 EPASEEPEADTDQARLS-----ADYKVELPLEDQVGDLEA-----SSEKCAPLA 379


```

Db 993 LSETERGEVEDEKENKEYSEKCVTSKQOEIETLYLKQITISEKEVELOHTRKDLK--LA 1049
Qy 380 TEVDEKMEAHQOEVAEYHVSTVEKTEBEOGGGEABGVVETGESLIPREKLAPEQEV 439
Db 1050 AE-----EFOQALVKQNMOTLQDKTNO-----IDLQAEISENOAI 1085
Qy 440 POEAPAEELKMSRMCVSGGDHQTLD-----LSPBEKTLPKHPEIGVISEVEMLSOERI 495
Db 1086 IQK-----LITNTDASDGSVALVKEVTIVISPPCTGSSEHMKPELEKILALEKEKE 1138
Qy 496 KVQGSPLKLPLSSSGLKKLSGKKQKGRGGGDEBEPGEYOHITHESPSADBOKSES--- 552
Db 1139 QLOKQLQALTSRKAILKKQAQEKERHLR-----BELKQKQXDYNRLQEQFPDQSENENI 1193
Qy 553 -----SASSPEPEBETTCLKGPPLAPDQGAEBGTT----- 584
Db 1194 GDQLRQLOIYVRESIDGKLPTDQOESCSSTPGLPEPLFKATEQHHTQPVLESNLCPDWP 1253
Qy 585 --SDGEKKRBEITTPWASPRKMWTPPKKVRPRPESDKEELEKVK--SATLSTDSSTVSEMQ 641
Db 1254 SHSEBASALQOGTSVAQIKAOI-----KEIEAEVLELELKVSSTTSSELTKKS 1300
Qy 642 DEVTKVEBQKPEBPKRRVDTSVSMEALICVSSSKRKARAKASSDDEGGPTLGGDSHRA 701
Db 1301 BEVFOLOQINKQGLEISLKTIVSHEAEVHAESLQOKL-----ESSOLOIAGLEHLR 1352
Qy 702 BEASDKRAGTDVAVASTQEOQOAGS---SSPEPAGSPEBEGVSTWESFRVLTPRKK 758
Db 1353 ELQPLDEL---QKLISSKEBVSYLSGOLSEKALATQIOETIIEQEDLICALHT--- 1405
Qy 759 SKSKLEBAEDSSVQOLSTET-----EPREBSWVSIKKRTPRRKKRAGKQBOAT 811
Db 1406 -QLEQAQAEHERIQOVLQELCEMKQKPELIGES-----RAKQOIRKQOALI 1454
Qy 812 EDSGVEINEDDPNVPAYVPLSEYNAVER-----EKMEAQGTLEPLQILGAVVSE 863
Db 1455 SRK---EALKKXSIQOEBELSLAR-GTIERLTSLADVESQVAKQKQDTVGRILALQE 1510
Qy 864 LSKTLVHTVSAVAVIDGTRAVTSVEBSPWISAVTEPLEHTAGAMPVEEVEETKDIJA 923
Db 1511 ERDKLITEMDRSLLENQSLSSCE--SLKLALGELTEDKELVKEL-----ESLSSKIA 1563
Qy 924 EETPVLTOTLPEBKAND-----MTSEVDFTSEAVTATETSEALRTEE----- 968
Db 1564 EST-----EMQKHELOKEYEYIILQSYENAVSNEARIGHVAVAEQOEYLGKLR 1615
Qy 969 VTEASGAETTDMVASVOLSPTTEBAPTVOEVESGVLDTEBEBRTOAILQAVADK 1028
Db 1616 STEANKKETEKOLQAEQOEMEMKEMKRF--AKSKQKILBLEBNDRILAEVHPADDT 1673
Qy 1029 VKESQVATQTVORTGSKALEKEVEBEDSEVLASE-----KEKDVMKGPQAEAGAH 1083
Db 1674 AKE-----CMETLLSSNMSMKELERVKMEYETLSKFFQSLMSEKDSISE-EVQD--LKH 1725
Qy 1084 LAQSGEQATPESLEVEBVTADVDHATQV---IKLOQLEQAVAPSSSTLLDSEIN 1140
Db 1726 QIEGVNSQOAMLEATEKDNQNTVTEBGTQSIPEGTEBODSLSMSTRPTCSVSAPSAXS- 1784
Qy 1141 GSTPLADBDTADGTQOD---ETIDSQDSKATAAVQSVTEBEAATAQOEESTPLPNVNP 1197
Db 1785 -ANPVSXKDFSSHDEINNYLOQIDQKERIAGLEBEEKQNNKFSOTLENEK--NTLLSQIS 1842
Qy 1198 AOEBHGEPEPRVLEPTQOELT-----AAAPVLAKTEVGOEGEDMWLDGKVKVE 1248
Db 1843 TKD--GE-----LKLQGEVTKMMLNQIOBELSRVTKLKEFAEBEDDLBERLMLNDL 1894
Qy 1249 QEVFVHSGPNSQKADVT-----YDSEVWGVAQCOEKESSTEVSLSLEGEMETDVE 1300
Db 1895 AEL---NGSIGNYCODVDIAQIKNELBSEMKNLKKCVSELBEKQOQLVKEKTKVESIR 1951
Qy 1301 KEKPRTKBEQVSEBEGQETAAPEHNGTYGKPVLLTDMSESENGKXLLGSLGSPSLPDDOX 1360
Db 1952 KEYLEKIQQAQKEPQNK-----SHAKQELQELLKEKQOEVKOL---QKDCIRYQEK 1998

```

```

Qy 1361 AGCIEVOYQSDITVTYQTAEVAEVEKVIETVVISETGESPECYGAH-----LLPAKESAT 1414
Db 1999 ISAEERTVKALEFVQTESQKLEITKEN--LAQAEHKKQAQELASFKVLLDQOSERA 2056
Qy 1415 GGHMTLQHAEDVPLGPSQA--ESIPITVPAPASTLHPDLOGEISASQERSEEBDKP 1472
Db 2057 -----RVLADNLKLELOSNKESVKSQMKQDE-----DLERLEBAEKHLKEKQM 2105
Qy 1473 DAGPDQKESSTALEKVL-----KAPELILESXNKIV 1507
Db 2106 QEKDALREKHEHLETTIGEIVTLNKKREVOQLQEWLSDTV 2148

RESULT 15
US-09-724-676A-55849
; Sequence 55849, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OR INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55849
; LENGTH: 2915
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-724-676A-55849

Query Match      5.0%, Score 400; DB 5; Length 2915;
Best Local Similarity 19.1%; Pred. No. 7.9e-15;
Matches 306; Conservative 304; Mismatches 663; Indels 330; Gaps 58;

Qy 43 ADADATVTLQKNGQLSGVNGVAGQDVHVOENQOGEVEEDVEDVQGRSEVDREKDR 102
Db 738 AEAERORRLDYES--QTARDNLLTTEO--IHSLSIEAKSKADVXI---EVLQNLBDVQ---- 787
Qy 103 VEEMANSTAVADITTKQOEBTSEI---EQIPASENNVEEMVOPAESQANDVGFKKVF 159
Db 788 -LQPEBOSTLIRLSQOLQNKSEVLBGAERVRHISKVEELSQLQSKLEI----- 839
Qy 160 FVGKFTYKQDNKESDVTQLLTVYKQEGEAGASVGAQDHOEBSVETAVESAKSESL 219
Db 840 -----TKMD-----QLLLEKKRDVETLQOTIEBKDOQVETISFSTEMKWLQNEB 884
Qy 220 KQSTFKQSGTLKQBSSTEIPLOAESDQAAEEAKDDEBEEKQEKPTSPSPSPVNSE 279
Db 885 KFSLGVEIKTLE-----QNLNLSFAEBAKKEQVEDNEVSSGLQNYDEMSPAQISKE 939
Qy 280 TTSSFKEPFTGMAWRKKTSPFKSK-----EDLETAERKEQBAEKVDEBEKEKT 331
Db 940 E-----LQHEFDLKKENQQRKXLOALLINRKELLQVRSRLEELANLKDESKKILP 992
Qy 332 EPASEQPAEDTQARIS-----ADYEKVELPLEDYQGDLEA-----SSEKCADLA 379
Db 993 LSETERGEVEDEKENKEYSEKCVTSKQOEIETLYLKQITISEKEVELOHTRKDLK--LA 1049
Qy 380 TEVDEKMEAHQOEVAEYHVSTVEKTEBEOGGGEABGVVETGESLIPREKLAPEQEV 439
Db 1050 AE-----EFOQALVKQNMOTLQDKTNO-----IDLQAEISENOAI 1085
Qy 440 POEAPAEELKMSRMCVSGGDHQTLD-----LSPBEKTLPKHPEIGVISEVEMLSOERI 495
Db 1086 IQK-----LITNTDASDGSVALVKEVTIVISPPCTGSSEHMKPELEKILALEKEKE 1138
Qy 496 KVQGSPLKLPLSSSGLKKLSGKKQKGRGGGDEBEPGEYOHITHESPSADBOKSES--- 552
Db 1139 QLOKQLQALTSRKAILKKQAQEKERHLR-----BELKQKQXDYNRLQEQFPDQSENENI 1193
Qy 553 -----SASSPEPEBETTCLKGPPLAPDQGAEBGTT----- 584

```

Db 1194 GDQLRQIQVRESIDGKLPSTDOQESCSSTPGLBEPFKATEQHHTQPVLESNLCPDWP 1253
QY 585 --SDGEKKREGITPWFASFKKMTPTKRVRRPSSDKEEELKVK-SATLSSTDSTVSEM 641
Db 1254 SHSDASALQGGTSVAQIKAQ-----KEIEAKVELELKVSTTSELTKS 1300
QY 642 DEVKTVGEEQKPEEPKRRVDTSVSWEALICVSSKKKARKASSDDDEGGPTLGGDSHRA 701
Db 1301 EEVFOQEQINKOGLETESLKTVSHEAEVHAESLQKL-----ESSOLQIAGLEHLR 1352
QY 702 EEASKDKEAGTDAVPASTQDQAQGS---SSPEPAGSPSEGGVSTWESFKRLVTPRKK 758
Db 1353 ELQPKDEL---QKLSKSEEDVYLSGQLSEKAALTKIQTIEQEDLIKALHT---- 1405
QY 759 SKSKLEKAEADSSVQELSTEI-----EPSREESVWSIKKFIPIGRKKRAGKQEOATV 811
Db 1406 -QLEMQAKEHDERIKQLQVELCEMKQKPEIGBES-----RAKQOIQRKLOAALI 1454
QY 812 EDSPVNEINDDPNVPAVPLSEYNAVER-----EKMEAQGNTELPOLLGNAVYVSEE 863
Db 1455 SRK---EALKENKSLQBELSLAR-GTIERLTKSLADVESQVSAQNKEKOTVIGRLALLQE 1510
QY 864 LSKTLVHTSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGAEAMPPVEEVTEKDIIA 923
Db 1511 ERDKLITEMDRSLLENQSLSSCE--SLKALBGLTEDEKELVKEI-----ESLKSKKIA 1563
QY 924 EETPVLTQTLPEGKDAHD-----WTVSEVDTSEAVTATSETSEALRTEE----- 968
Db 1564 EST-----EWOEKHELQKEVEILLQSVENVSNEAERIQHVVAEAVRQEKQELYGKLR 1615
QY 969 VTASGAEEITDMVASVSQLTSDPTTEATPVQEVESGVLDTTEERERQTAILOAVADK 1028
Db 1616 STEANKKETQKQAEAEQEMEEMKEKMKRF--AKSKOOKILELEEEENDRLRAEVHPAGDT 1673
QY 1029 VKESQVPAQTQVORTGSKALEKVEEVEDSEVLASE-----KEQVMPKGPVQEAAGAH 1083
Db 1674 AKE-----CMETLLSSNASMKEELERVMEYETLSKKFQSLMSEKDSLSE-EVQD--LKH 1725
QY 1084 LAQGETQOATPESLEVPVADVHDVATCQV---IKLQQLMEQAVAPESSETLTDSETN 1140
Db 1726 QIEGNVSKQANLEATEKHDQNTNVTEBGTQSI PGTEETEEDQSLSMSTRPTCSSESVPSAKS- 1784
QY 1141 GSTPLADSDTADGTQDP---ETDSDQSKATAAVRQSVTTEEBAAATAQKEEPSTLPNNVP 1197
Db 1785 -ANPAVSKDFSSHDEINNYLQOIDQLKRIAGLEEEERQKNEKFSQTLNEK-NTLLSQIS 1842
QY 1198 AQEBHGEPGRDVLPTQBELT-----AAAVPVLAKEVQGEVQDVLGDKVKKEE 1248
Db 1843 TKD--GE-----LKMLOEVTWMLNNOIQBELSRVTKLKTABEEKODLEERLMNQL 1894
QY 1249 QEVFVHSGPNSOKAADVT-----YDSEVMGVAGQEKESTEVQSLSLEEGEMETDVE 1300
Db 1895 AEL---NGSIGNYCQDVTDAIKNELLESEMKNLKKCVSELEEEKQLVKETKVESEIR 1951
QY 1301 KEKRETPQEVQSEGEQETAPEHEGTYGKPVLTLDMPSSERKALGSLGSPSLPDQDK 1360
Db 1952 KEYLEKTOGAQKEPGNK-----SHAKELQELLKQKEQVKQL-----QKDCIRYQEK 1998
QY 1361 AGCIEVOVQSLDVTVTOTABAVERKVIETVWISETGESPECVGAH-----LLPAEKSSAT 1414
Db 1999 ISALERTVKALEFVQTESQKDLITKEN--LAQAVEHRKKAQELASFVLLDDTQSEAA 2056
QY 1415 GGHWTLQHAEDTVPLGPESQA--ESIPITVPAPESTLHPDLOGEISASQERSEEDKP 1472
Db 2057 -----RVLADNLKLLKELQSNKESVKSQMKQDE-----DLERRLEQAEKHLKEKKNM 2105
QY 1473 DAGPDADGKESIAIEKVL-----KAPPEILELESKSNKIV 1507
Db 2106 QEKLDALRREKRVHLEETIGIEIQTVLNKKQKEVQOQLENLDSTV 2148

THIS PAGE BLANK (USPTO)

GenCore version 5.1.13
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 00:28:41 ; Search time 6461 Seconds
(without alignments)
6210.712 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSSTQRSEQPGASDT.....AWAQRKCLPRLQLKAPVSK 1596

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues
Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US0902432/runat_12122002_141453_25861/app_query.fasta_1.1735
-DB=Pending Patents NA_Main -OFMT=fastap -SUFFIX=nrpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0902432 @CGN 1 1 5452 @runat_12122002_141453_25861 -NCPU=6 -ICPU=3
-NO XLPXY -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA_Main:
1: /cgn2_6/ptodata/2/pna/PCTUS COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081 COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082 COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083 COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084 COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085 COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087 COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088 COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089 COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090 COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091 COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092 COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093 COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US094 COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US095A COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US095B COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095C COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095D COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US096A COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US096B COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096C COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096D COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096E COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US097A COMB.seq.*

29: /cgn2_6/ptodata/2/pna/US097B COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097C COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US098A COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US098B COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098C COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US099A COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US099B COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US099C COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US099D COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US100A COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*
44: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6023 COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6026 COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8073	100.0	5200	13	US-08-978-277A-3
2	8073	100.0	6160	1	PCT-US97-06830-3
3	8073	100.0	6160	10	US-08-665-401-3
4	8073	100.0	6160	34	US-03-902-432-3
5	8030	99.5	5236	41	PCT-US02-16173A-2063
6	8030	99.5	5236	41	US-10-191-803-279

Sequence 12210, A
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 4018, Ap
Sequence 10143, A
Sequence 517, App
Sequence 6364, Ap
Sequence 9117, A
Sequence 10117, A
Sequence 10175, A
Sequence 8758, Ap
Sequence 2000, Ap
Sequence 11667, A
Sequence 6654, Ap
Sequence 5679, Ap
Sequence 3215, Ap
Sequence 3428, Ap
Sequence 4599, Ap
Sequence 2245, Ap
Sequence 701, Ap
Sequence 2407, Ap
Sequence 1797, Ap
Sequence 5392, Ap
Sequence 5120, Ap
Sequence 30580, A
Sequence 2, Appli
Sequence 58, Appl
Sequence 1, Appli
Sequence 3439, Ap
Sequence 58, Appl
Sequence 21009, A
Sequence 97, Appl
Sequence 30584, A
Sequence 282, App
Sequence 24882, A
Sequence 10145, A
Sequence 7019, Ap

RESULT 1
US-08-978-277A-3

	TELECOMMUNICATION INFORMATION:
	TELEPHONE: 212-408-2558
	TELEFAX: 212-765-2519
	TELEX:
	INFORMATION FOR SEQ ID NO: 3:
	SEQUENCE CHARACTERISTICS:
	LENGTH: 5200 base pairs
	TYPE: nucleic acid
	STRANDEDNESS: single
	TOPOLOGY: linear
	MOLECULE TYPE: cDNA
	HYPOTHETICAL: NO
	ANTI-SENSE: NO
	FRAGMENT TYPE:
	ORIGINAL SOURCE:
	US-08-978-277A-3
	Alignment Scores:
	Pred. No.: 0
	Score: 8073.00
	Percent Similarity: 100.00%
	Best Local Similarity: 100.00%
	Query Match: 100.00%
	DB: 13
	Gaps: 0
	Length: 5200
	Matches: 156
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0
	US-09-902-432-4 (1-1596) x US-08-978-277A-3 (1-5200)
QY	1 MetGlyAlaGlySerSerThrGlnArgSerProGlnInProAlaGlySerAspThr
Db*	1 ATGGGCGCAGCGCAATTCACCGAGCAGCGAGCCCGACGACCGCGGGAGGACAGC
QY	21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGlnAlaSerGlyAlaAlaGly
Db	61 CCGAGCGAGCTGGTGGCTCAGTGGCCATGGCCCGCAGCTGAAGCTCTGGAGCAGCTGGA
QY	41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer
Db	121 GACCCCGCCAGCGGGGACCCCGCCACCAAGCTCCCAAGAAAGATGGCCAGCTCTTCT
QY	61 ValAsnGlyValAlaGlnGlnGlyAspValHisValGlnGlnGlnGlnGlnGlnGln
Db	181 GTCAACGGCGTAGCTGAACAAGAGATGTCATGTCACAGGAGAAACAGAGAGGGCAG
QY	81 GluGlnGluValValAspGluAspValGlyGlnArgGlnSerGluAspValAaGlnLys
Db	241 GAGAAAGAAAGTCGTGTATGAGATGTTGGACAGGACAGACAGTCAGAAAGATCGAGGAAAA
QY	101 AspArgValGlnGlnMetAlaAlaAsnSerThrAlaValGlnAspLysLysAspGly
Db	301 GACCGAGTTGAAGAAATGGCGGCCAATCTCCACACACTGTGTGAAGATATCAAGAGATGG
QY	121 GlnGlnGlnThrSerGlnLysLeuGlnLysLeuProAlaSerGluAsnValGlnGln
Db	361 CAGAGGAGAGATCAGAAATATTTGAACAGATCCCTCTCTCAGAAAAATATGTGGAGAA
QY	141 MetValGlnProAlaGlnSerGlnAlaAsnAspValGlyPheLysValPheLysPhe
Db	421 ATGTACAGCTCTGTGATGCCCAAGGCTAAATGATGTTGGCTTCAAGAAAGATTTAAATTT
QY	161 ValGlyPheLysPheThrValLysAspLysAsnGlnLysSerSerAspThrValGlnLeu
Db	481 GTTGATTTTAATTCACGGGTGAAGAGATTAATAATAAAGTCAGATCTGTCACATA
QY	181 LeuThrValLysLysAspGlnGlyGlnGlyAlaGlnAlaSerValGlyAlaGlyAspHis
Db	541 CTCACCTGTCAAGAAAGATGAAGGCGGAGGCGGACCAACCTCTGTGGAGCTGAGAGACAC
QY	201 GlnGlnProSerValGlnThrAlaValGlyGlnSerLysSerGlnSerGlnLeuLys
Db	601 CAGAGCCCAAGTGTGAGAACTGCGGTGGAGAGTCACATCCAAAGAAATGAGCTGAG
QY	221 GlnSerThrGlnLysGlnGlnGlnGlyThrLeuLysGlnGlnGlnSerSerThrGlnLysPro

Db 661 CAATCCACAGAGCAAGCAAGCCCTCAGCAAGCAAGACAGAGCAGCACAGAAATCCCC 720
 Qy 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLys 260
 Db 721 CTTCAAGCCGAATCTGATCAAGCGCTGAGGAAGAGCCAAAGATGAAGAGAGAAAA 780
 Qy 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAenSerGluThr 280
 Db 781 CAAGAGAAAGAGGCCCAAGTCCCCAGAAATCCCGAGCAGCCAGTCAACAGTGAGACA 840
 Qy 281 ThrSerSerPheLysLysPheThrHisGlyTyrAlaGlyTyrArgLysLysThrSer 300
 Db 841 ACATCTCTTCAAGAGTTCTTCACTACGGTTGGCCCGCTGGCCCAAGAGACAGC 900
 Qy 301 PheLysLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
 Db 901 TTCAAGAAATCAAAAGAGGATGATCTGAAACTGCCGAGAAAGAGAGAGAGGCA 960
 Qy 321 GluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGluGlnGluPro 340
 Db 961 GAAAAAGTAGACGAGGAAGAAAGAAAGACAGAGCCCTCGGAGGAGCAGAGCCG 1020
 Qy 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeu 360
 Db 1021 GCAGAGACACACAGCCAGGCTGTTGACGACTACAGAGAGGTGGAGCTGCCTTTG 1080
 Qy 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThr 380
 Db 1081 GAAGACCAGGTGGTGACCTGGAGGCATCGTCAGAGAGAGAGTGTGCTCTTTGGCAAG 1140
 Qy 381 GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
 Db 1141 GAAGTGTGTTGATGAGAGATGGAAGCCCAACCAAGATGTTGACAGAGGTCCACGTGAGC 1200
 Qy 401 ThrValGluLysThrGluGluGlnGlyGlyGlyGlyGluAlaGluGlyGlyValVal 420
 Db 1201 ACCGTGAGAAAGACAGAGGAGGAGCAGGAGAGGAGGAGGAGGCTGAAGGGGGCGTGGT 1260
 Qy 421 ValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValPro 440
 Db 1261 GTAGAGGAACAGAGAAATCCTTGGCCCTCGAGAACTGGCTGAGCCCAAGAGGTCCCC 1320
 Qy 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
 Db 1321 CAGGAAGCTGAGCGTCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAGGA 1380
 Qy 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
 Db 1381 GACCACACTCACTGACAGACCTAAGTCTTGAAGAGAGAGCGTGGCCCAACACCCAGAA 1440
 Qy 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500
 Db 1441 GGCATTGTAGTGGGTGGAGATGCTGTCTCTCAGAAAGAATCAAGGTACAGGAGT 1500
 Qy 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
 Db 1501 CCCTTGAAGAAACTCTTTCAGTAGCTCAGGCTTAAAGAAGCTGTCTGGGAAGAGCAGAA 1560
 Qy 521 GlyIleArgGlyGlyGlyAspGluGluProGlyGluThrGlnHisIleHisThrGlu 540
 Db 1561 GGGAAACGAGAGGTGGGGAGAGCAGAGAGCCTGGAGAAATACCAACATTCACCCGAA 1620
 Qy 541 SerProGluSerAlaAspGluGlnLysGlyLysSerAlaSerSerProGluGluPro 560
 Db 1621 TCCCCAGAGAGTGTGATGACAGAGAGAGAGCTTCTGCTGCTCCCCCGAGAGCCT 1680
 Qy 561 GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
 Db 1681 GAGGAGACCAGCTGTCTGGAGAAAGGCGCTGGAAGCACCACCCAGAGATGGGAAGCTGAG 1740
 Qy 581 GluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProThrAlaSerPhe 600
 Db 1741 GAAGGAACTACTTCCGATGGAGAGAGAGAGAGGATCACTCCCTGGGCATCCTTC 1800

Qy 601 LysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGluGlu 620
 Db 1801 AAAAAGATGTTGACACCCCAAGAAACCGGTCCGAAGACCTTCTGAGAGTGACAAGAGGAA 1860
 Qy 621 GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
 Db 1861 GAGCTGGAGAAAGTCAAGAGCGCCACTTGTCTCCACTGATAGCACAGTGTGAGAAATG 1920
 Qy 641 GlnAspGluValLysThrValGlyGluGlnLysProGluProLysArgVal 660
 Db 1921 CAAGATGAAGTCAAAACTGTGTGTGAGGAACAAAGCCAGAGAAACCAAGCGTAGGGTG 1980
 Qy 661 AspThrSerValSerTrpGluAlaLeuLysCysValGlySerSerLysLysArgAlaArg 680
 Db 1981 GATPACTTCAGTGTCTTGGGAAGCACATGATTTGTGTCCGATCATCCCAAGAGAGCAAGG 2040
 Qy 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
 Db 2041 AAGGCATCTCTTCAGATGATGAAGAGGGCCCAAGGACACTGGGAGGGGACAGTCAAGA 2100
 Qy 701 AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
 Db 2101 GCAGAGAGGCCAGCAAGCAAGAAAGCCGACAGAGCGCTGTTCTGCCAGACCCAG 2160
 Qy 721 GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly 740
 Db 2161 GAGCAGGACCAAGCGCAAGAGTTCCTCACCCAGCAGCGGGAAGCCCTCCGAAGGG 2220
 Qy 741 GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
 Db 2221 GAAGGTGTCTCCACTTGGGAGTCATTTAAAGATTAAGTCACTCCAAGAAAAATCCAAG 2280
 Qy 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
 Db 2281 TCAAACTGGAAGAGAAACCCGAGACTCTAGTGTAGACAGCTTGCTCACTGAGATCGAA 2340
 Qy 781 ProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArgLysLys 800
 Db 2341 CCGAGTAGAGAGAAATCTTGGGTTCATTAAGAAATTCATCCCGGCGCGGAGAA 2400
 Qy 801 ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAen 820
 Db 2401 AGGCGAGAGCGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2460
 Qy 821 GluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaValGluArg 840
 Db 2461 GAGGACGACCTTAATGTCCAGCCGCTGCTGTCTGAGTATTAATGAGTGGAGAGG 2520
 Qy 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal 860
 Db 2521 GAGAAGATGGAAGCCAGGGGAATACGGAGCTGCCAGCTGCTGGGGCTGTGTACGTG 2580
 Qy 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
 Db 2581 TCCGAGGAGCTCAGTAAAGACTCTGTGTCCACACTGTGAGTGTGAGTGTGAGTGTGAG 2640
 Qy 881 ArgAlaValThrSerValGluArgSerProSerTrpIleSerAlaSerValThrGlu 900
 Db 2641 AGGGCAGTCACCAAGTTCGAAGAGCGGTCTCTCTGTGTGATATCCCGTCCCGTAAACAGAA 2700
 Qy 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
 Db 2701 CCTCTTGAACACACAGCGGAGAGCCATGCCACTGTTGAAGAGGTCACTGAAAAGAC 2760
 Qy 921 IleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
 Db 2761 ATCATTGCAAGAGAAACTCTGTGTGCTCAACCCAGACGTTACCAGAGGTAAAGATGCCAT 2820
 Qy 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
 Db 2821 GAGGACATGTTCAACAGTGAAGTGGATTTCACTCAAGAGCTGTGACAGCCACAGAGACC 2880

[illegible]

Dd	3961	GCTCCTGAGCATGGAAGAACTTACGGGAAAGCCAGTCTTGACCTTGACATGCGCCAGCTCA	4020
Qy	1341	GUUARGGLYUYSALALEUGLISERLEUGLIGLYSERPROSERLEUPROASPLINAPLS	1360
Dd	4021	GAGAGGGGGAAAGCAGCTGGGAAACCTTTGAGAGAGGCCCTTCTCTCCAGACCAAGACAA	4080
Qy	1361	ALAAGLYCSYILEGLUVALGINVALGINSERLEUASPTHTRVALTHGINTRVALIGU	1380
Dd	4081	GCAGGTGTCATAGAGGTTCAAAGTTCAAACCTTGACACAAACAGTCACTCAAAACAGCAAA	4140
Qy	1381	ALAVALGULYUYSVALILEGUTRTHRVALLIISERGIUTRGLYUYSERPROGLUCYS	1400
Dd	4141	GCTGGGAAAAGGTCATAGAAACGGTTGGATTTCAGACACAGGTGAAGGTCAGAGGT	4200
Qy	1401	VALGILYLAHISLEULEUPROALIGULYSSERSEALATHGLYVHISITRPTHREU	1420
Dd	4201	GTAGGTGCACACTTATTATCCAGCTGGAAGTCTCTGCAACGGGTGGCCATCGACTCTT	4260
Qy	1421	GINHISALAGUASPTHRVALLPROLEUGLYPROGLUSERGINALIGUSERILEPROILE	1440
Dd	4261	CAGATGCAGAGGACACGGTAAACCCCTGGGGCCCTGAGTCTCAGGCAGAAATCCATCCAAATC	4320
Qy	1441	ILEVALTHIRPROALAPROGLUSERTHIRLEUHSIROASPLEUGINGLYVLIISERILA	1460
Dd	4321	ATAATGACTCCTGCTCTGTAAACACCTCAATCTCGACCTTCAAGGAAATAAGCGCA	4380
Qy	1461	SERGINRUGGUARGSERGIUGIUGUASPLYSPROASPLAIGLYPROASPLAASPLY	1480
Dd	4381	TCCAGAGAGACGCATATAGAGAAAGGACACACCGACGATCTGGTCTGATGCTGACGGC	4440
Qy	1481	LYSGUGSERTHIRALALEGLUYVALLEULYSAIAGUPROGLUILEUGLUUGLU	1500
Dd	4441	AAGGAGGTACAGCAATCGAAAAGTCTTCAAGGCTGAACCTGAGATCTGGAACCTGAG	4500
Qy	1501	SERYSSERSENLYSILEVALILEUAENVALLILEGINTHRALAVALASGINPHEALARG	1520
Dd	4501	AGTAAGAGCAACAAGATTGGCTGAAACGTCATTCAAGACCGCGTTGACAGTTCGCACGT	4560
Qy	1521	THRGUITHRALAPROGIUTRTHIRHISAIARYAASPERGINTHRGIVALPROALACYSARG	1540
Dd	4561	ACAGAAACAGCCCCCGAAACTCATGCTTATGATTACACAGACCCAGATTCTTGATCGACG	4620
Qy	1541	LEUASPERARGIUPROASNARGCYSTRPTHLYSMETLYASPLAALYSMETLYSHIS	1560
Dd	4621	CTTGACAGCAGGGAGCCCAACAGATCTGAGCAAAAATGAAGATGCCAAGTGAACAC	4680
Qy	1561	PROVALPROGINPROARGIUNASPLEUGINGVALLEUTHRVALLILEUGLUVALTPALAGIN	1580
Dd	4681	CCAGTGGCCGACGCCACAGAGGACTTGCAAGTCTCGAACCGTTCGTGAAGCATGGGCTCAG	4740
Qy	1581	PROARGLYSCYLEUPROARGLEUGINLEULYSAIAPROVALISERYLS	1596
Dd	4741	CCTGGAAATGCTTGCCGCGCTTGCAAGTTGAAAAGCGCGGTGTCAAG	4788
RESULT 2			
PCT-US97-06830-3			
Sequence 3, Application PC/TUS9706830			
GENERAL INFORMATION:			
APPLICANT: Gelman, Irwin H.			
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE			
NUMBER OF SEQUENCES: 20			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond			
STREET: 30 Rockefeller Plaza			
CITY: New York			
STATE: NY			
COUNTRY: USA			
ZIP: 10112-0228			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Diskette			
COMPUTER: IBM Compatible			
OPERATING SYSTEM: DOS			

```

1 RESULT 2
2 PCT-US97-06830-3
3 ; Sequence 3 Application PC/TUS9706830
4 ; GENERAL INFORMATION:
5 ; APPLICANT: Gelman, Irwin H.
6 ; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
7 ; NUMBER OF SEQUENCES: 20
8 ; CORRESPONDENCE ADDRESS:
9 ; ADDRESSEE: Brundagh, Graves, Donohue & Raymond
10 ; STREET: 30 Rockefeller Plaza
11 ; CITY: New York
12 ; STATE: NY
13 ; COUNTRY: USA
14 ; ZIP: 10112-0228
15 ; COMPUTER READABLE FORM:
16 ; MEDIUM TYPE: Diskette
17 ; COMPUTER: IBM Compatible
18 ; OPERATING SYSTEM: DOS
19 ;

```


; SOFTWARE: FastSeq Version 1.5
 ; CURRENT APPLICATION DATA: PCT/US97/06830
 ; FILING DATE: 18-JUN-1996
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA: 08/665,401
 ; APPLICATION NUMBER: 08/665,401
 ; FILING DATE: 18-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Clark, Richard S
 ; REGISTRATION NUMBER: 26,154
 ; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-408-2558
 ; TELEFAX: 212-765-2519
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6160 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE:
 ; ORIGINAL SOURCE:
 ; PCT-US97-06830-3

Alignment Scores:
 Pred. No.: 0 Length: 6160
 Score: 8073.00 Matches: 1596
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-09-902-432-4 (1-1596) x PCT-US97-06830-3 (1-6160)

QY 1 MetGlyAlaGlySerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
 DB 33 ATGGCGCAGGAGTTCACCGAGCAGCGGAGCCCGGAGGAGCGGCGGAGGAGCAGC 92
 QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
 DB 93 CCGAGCGAGTGGTGTCTCAGTGGCCATGGGCGGCGGAGTGAAGCTCGGAGCAGCTGGA 152
 QY 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
 DB 153 GACCCCGCGAGCGGAGCCCGCCACCAAGCTCCACAGAGAGATGGCCAGCTGCTTCT 212
 QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGlnAsnGlnGlyGln 80
 DB 213 GTNACCGGTAGCTGACACAGAGAGATGTCATGTCACAGAGGAAACACAGAGGCGGAG 272
 QY 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
 DB 273 GAGGAGAAGTCGTTGATGAGGATGTTGGACAGCGAGAGTCCAGAGATGTCAGAGAAAAA 332
 QY 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
 DB 333 GACCGAGTTGAAGAAATGGCGGCCCACTCCACAGCTGTTGAAGATATACAAAAGGATGG 392
 QY 121 GlnGluGluThrSerGluIleGluGlnIleProAlaSerGluAsnAsnValGluGlu 140
 DB 393 CAGAGAGAGACATCAGAAATTAATTAACAGATCCCTGCTTCAGAAACCAATGTGGAAGA 452
 QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPhePhe 160
 DB 453 ATGGTACAGCTGCTGAGTCCAGGCTAATGATGTTGGCTTCAAGAAAGTATTAAATTT 512
 QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180

DB 513 GTTGGTTTTAAATTACGGTGAAGAAGGATAAAAAATGAAAAGTCAGATACTGTCCAACTA 572
 QY 181 LeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
 DB 573 CTCACTGTCAAGAAGGATGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 632
 QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
 DB 633 CAGGAGCCCAAGTGTGAGAGCTGCGGTGGAGAGTCAAGATCCAAAGAAAGTGAAGTGA 692
 QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIlePro 240
 DB 693 CAATCCACAGAGAGACAGAGGAGCCTCTGAACAGCAAGACAGACAGACAGCAAGAAATCCC 752
 QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGluLys 260
 DB 753 CTTCAAGCGGAATCTGATCAAGCGGTGAGGAAGAACCAAGATGAAGGAGAAGAAAA 812
 QY 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThr 280
 DB 813 CAAGAGAAAGAGAGCCCAAGTCCCGAGATCCCGAGCAGCCCAAGTCAACAGTGAGACA 872
 QY 281 ThrSerSerPheLysLysPhePheThrHisGlyTyrAlaGlyTyrArgLysLysThrSer 300
 DB 873 ACATCTCTTCAAGAGAGTCTTCACTCACGGTTGGCCCGGCTGGCCCAAGAGACAGC 932
 QY 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
 DB 933 TTCAAGAAATCAAAAGAGGATGATCTGGAAACTGCCGAGAAGAGAGAGGAGCAAGAGCA 992
 QY 321 GluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGluGlnGluPro 340
 DB 993 GAAAAAGTAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1052
 QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeu 360
 DB 1053 GCAGAGACACAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1112
 QY 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThr 380
 DB 1113 GAAGACCAGGTGGTGCACCTGGAGGAGCATCGTCAGAGGAGAGAGTGTGCTCTTTGGCA 1172
 QY 381 GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
 DB 1173 GAAGTGTTCATGAGAAGATGGAAGCCCAAGAGAGATTTCTTCAGAGAGTCCAGCTGAG 1232
 QY 401 ThrValGluLysThrGluGluGlnGlnGlyGlyGlyGluAlaGluGlyValVal 420
 DB 1233 ACCGTGGAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1292
 QY 421 ValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValPro 440
 DB 1293 GTAGAGGAGACAGAGAGATCTTTGCCCTGTAGAGAACTGGCTGAGCCCGAGGAGTCCC 1352
 QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
 DB 1353 CAGGAAGCTGAGCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTCTGGAGGA 1412
 QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
 DB 1413 GACCCACATCACTGACAGACCTTAAGTCTTGAAGAGAGAGCGCTGCCCAACACCCAGAA 1472
 QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500
 DB 1473 GGCAATGTCTAGTGGTGGAGATGCTGTCTCTCAGGAAAGAAATCAAGGTACAGGGAGT 1532
 QY 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
 DB 1533 CCCTTGAAGAAATCTTTCAGTAGCTCAGGCTTAAGAGAGCTGTCTGGGAAGAGAGCAAG 1592
 QY 521 GlyLysArgGlyGlyGlyAspGluGluProGlyGluTyrGlnHisIleHisThrGlu 540
 DB 1593 GGGAAACGAGGAGGTGGGGAGAGCAGAGAGCTGGAGATACCAACACATTCACCCAGAA 1652

QY	541	SePrProGluSerAlaapbGluInuLysGlyGluSerSerAlaSerSerProGluInuPro	560
Db	1653	TCCCAGAGAGATGCTGATGTAGACGAAGGGGAGAGAGCTTCGCGTCGCCCCAGAGAGACT	1712
QY	561	GluGluThrThrCysLeuGluGluLysGlyProLeuGluAlaProGluAapbGlyGluAlaGlu	580
Db	1713	GAGGAGACCACGTTCTTGGAGAAAGGGCCGCTGGAGACACCCAGATGGGAACTGAG	1772
QY	581	GluGlyThrThrSerAapGlyGluLysLysArgGluGlyIleThrProTPAlaSerPhe	600
Db	1773	GAGGAACCTACTTCCGATGGAGAGAAAGAGAGAGAGGATCACTCCCTGGGACTCTTC	1832
QY	601	LysLysMetValThrProLysLysArgValArgArgProSerGluSerAplyGluGlu	620
Db	1833	AAAAAGAGGTGACACCCCAAGAAACGGGCTCCGAAGACCTTCGAGGTGACAAAGAGAA	1892
QY	621	GluLeuGluLysValLysSerAlaThrLeuSerSerThrApsSerThrValSerGluMet	640
Db	1893	GAGCTGGAGAAAGGTCAAGAGGCGCACTTCTCTCCACTGATGCACAAGTGTCAAGAAATG	1952
QY	641	GluAapGluValLysThrValGlyGluGluGluLysProGluInuProLysArgArgVal	660
Db	1953	CAAGTGAAGTCAAAACTGTTGGTAGAGAAACAAACACAGAGSAAACAAAGCTAGGGTG	2012
QY	661	AspThrSerValSerTPGluAlaLeuIleCysValGlySerSerLysArgAlaArg	680
Db	2013	GATACCTCAGAGTCTTGGGAAGCACTGATTGTGTGGATCATCTCCAAAGAGAGCAAG	2072
QY	681	LysAlaSerSerSerAapAapGluGlyGlyProArgThrLeuGlyGlyLysPheHisArg	700
Db	2073	AAGGATCTCTTCAAGATGATGAAGAGGGGCGCAAGACACTGGAGGGGACGTACAGA	2132
QY	701	AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln	720
Db	2133	GCAAGAGAGGCCACCAAAAGCAAAAGAACCGGGAACAGACCGTCTTCTCCAGACACCG	2192
QY	721	GluGluAapGluAlaGluGlySerSerSerProGluProAlaGlySerProSerGluGly	740
Db	2193	GAGCAGACCAAGCGCAAGAAAGTTCTCTCACCGAGCCAGCGGGAAGCCCTTCCCAAGGG	2252
QY	741	GluGlyValSerThrTPGluSerPheLysArgLeuValThrProArgLysLysSerLys	760
Db	2253	GAAAGTGTCTCCACTTGGGAATCTTTAAAGATTTAGTCACTCCAAAGAAAAAATCCAG	2312
QY	761	SerLysLeuGluGluLysAlaGluAapSerSerValGluInuSerThrGluLeuGlu	780
Db	2313	TCNAACCTGGAAAGAAAGCCGAAGACTCTAGGTGAGACAGTGTCCACTAGATCGAA	2372
QY	781	ProSerArgGluGluSerTPValSerIleLysLysPheIleProGlyArgArgLysLys	800
Db	2373	CCGAGTAGAAGAAAGATCTTGGGTTTTCATTAAAGAAATTCATCCCGGAGCGGGAAGAA	2432
QY	801	ArgAlaAapGlyLysGluGluGluAlaThrValGluLysPserGlyProValGluLeuLeu	820
Db	2433	AGGGCAGACGGGAAGCAAGAACCAAGCACTGATGAAGACTCAGGGCCAGATGAGATTAAT	2492
QY	821	GluAapAapProAenValProAlaValValProLeuSerGlyLysAsnAlaValGluArg	840
Db	2493	GAGGACGACCTTAATGTCCAGCCGTGTGCTCTGTCTGAGTAAATACAGTGGAGAGG	2552
QY	841	GluLysMetGluAlaGluGlyAsnThrGluLeuProGluInuLeuGluValAlaValTyrVal	860
Db	2553	GAGAAAGTAGAAGCCCAAGGGGAATACAGGACTGCCAGCTGTGGGGCTGTGTAACGTG	2612
QY	861	SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAapGlyThr	880
Db	2613	TCCCGAGAGCTCAAGTAAGACTCTGTGTCCACACTGTGAGTGTCCCACTCAATATGAGGACC	2672
QY	881	ArgAlaValThrSerValGluGluLysArgSerProSerTPLeuSerLysSerValThrGlu	900
Db	2673	AGGGCAGTCCACAGTGTCCGAAGACGGGTGTCTTTCGGGATATCCGCTTCCGTAAACAGA	2732

QY	901	ProLeuGluHisThrLysArgLysLysIleMetProProValGluGluValAlaThrGluLysAsp	920
Db	2733	CCTCTTAAACACACACGGGAGAAAGCAAGCCACTGTTGAAGAGCTCACTGAAGAAAGAC	2792
QY	921	IleIleLeuGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis	940
Db	2793	ATCATTCGAAAGAAACCTCTGTGCTCAACCCAGACGCTTACCAAGAGGTGAAGATGCCCCAT	2852
QY	941	AspAspMetValThrSerGluValAspPheThrSerGluAlaValAlaThrAlaThrGluThr	960
Db	2853	GACGACATGCTCACAGTGAAGTGAATTCACCTCAAGAAAGCTGTGACACGACAGAAACC	2912
QY	961	SerGluAlaLeuArgThrGluGluValAlaThrGluAlaSerGlyAlaGluGluThrThrAsp	980
Db	2913	TCAAGAGCTCTCCGTACTGAAGAGTTACCAGGAAGATGGGGGCGGGAAGAACCAACAGAC	2972
QY	981	MetValSerAlaValSerGluLeuThrAspSerProAspThrThrGluAlaAlaThrPro	1000
Db	2973	ATGGTGTCGCCAGATTTCACGCTGACTGACTCCCAAGACACCAAGAGAAAGCCACC	3032
QY	1001	ValGluGluValGluSerGlyValLeuAspThrGluGluGluValArgGluThrGluAla	1020
Db	3033	GTTCCAGAGGAGAGAGTGGTGCTAGATACAAAGAAAGAGAGAGCCGACAGCCAGGCC	3092
QY	1021	IleLeuGluAlaValAlaAspLysValLysGluGluSerGluValProAlaThrGlnThr	1040
Db	3093	ATTCCTCAACCCGTTGTGACGACAAAGTGAAGAGAGGATCCAGGTGCCGCAACCAACT	3152
QY	1041	ValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGlu	1060
Db	3153	GTCGAGAAACGGGGTCTCAAAAGCATCTGGAGAAAGTTGAGAGGTTAGAGAGAGACTCCGAA	3212
QY	1061	ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGluGluAlaGly	1080
Db	3213	GTGCTGGCTTCGGAGAAAGAAAGAGACGTTATGCCAAAGAGCCGTGCAAGAAAGCTGGA	3272
QY	1081	AlaGluHisLeuAlaGlnGlySerGluThrGlyGluAlaThrProGluSerLeuGluVal	1100
Db	3273	GCTGAGCATCTGTGACAGGGCTCTGTGACCTGACAGCGCTACTCCAGAGAGCTTTGAAGATT	3332
QY	1101	ProGluValThrLysAspValAspPheValAlaThrCysGlnValIleLysLeuGlnGln	1120
Db	3333	CCTAAAGCACGGAGATGTAAGCCATGTGCGACAGTGCACAGTTATCAAGCTCCAGAG	3392
QY	1121	LeuMetGluGluAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn	1140
Db	3393	CTGATGGAACAGGCGCGTGGCCCTGATGATCATCGAAACCTTGACAGCAAGTGAACAAAT	3452
QY	1141	GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnAspGluThrIle	1160
Db	3453	GGAAGCACTCCCTTAGACAGATTCAGACACTGCAAGATGGGACACAGCAAGATGAAGAACTAT	3512
QY	1161	AspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGluValThrGluGluGlu	1180
Db	3513	GACGACGAGACAGTAAAGCACTGCACTGTGACAGCATCAAGTCAACAGAAAGAG	3572
QY	1181	AlaAlaThrAlaGluLysGluGluProSerThrLeuProAsnAspValProAlaGlnGlu	1200
Db	3573	GCGGCTACTGTCAGAAAGAGAGCTTGACACTACCTAAATATGTTCCAGGCCAGGAA	3632
QY	1201	GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAla	1220
Db	3633	GAACTATGGGGAAGAACAGAGAAAGATGTTCTTGAACCTACACACAGACAGACTTACTCT	3692
QY	1221	AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyValAspThrLeu	1240
Db	3693	GCAAGCCGCGCCGTTCTGGCAAAAGACTGAGGTGGGTCAAGAGGGTGAAGTTGACGTGGT	3752
QY	1241	AspGlyGluLysValLysGluGluGlnGluValPheValHisSerGlyProAsnSerGln	1260
Db	3753	GATGGAGAAAGCAAGCAAGAGAAACAGAGAGGTGTTGTACACTCTGGACCCAAACGTTAA	3812
QY	1261	LysAlaIleAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLys	1280

```

Db 3813 AAGCTCTGATGACATATGACAGTGAAGTATGGAGTGGCCGGTGTCCAGAAAG 3872
Qy 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluGluMetGluThrAspValGlu 1300
Db 3873 GAGAGTACTGAAGTGCAGAGCTTAGCTGGAGGAGGAGATGGAATCTGACGTGAA 3932
Qy 1301 LysGluLysArgGluThrLysProGluValSerGluGluGluGluGluThrAla 1320
Db 3933 AAGGAGAAAGGGAGACAAAGCCAGACAGTGAAGTGAAGAGGTGACGAGAAACAGCC 3992
Qy 1321 AlaProGluHisGluGluThrTyrgLysProValLeuThrLeuAspMetProSerSer 1340
Db 3993 GCTCTGAGCATGAGGAACTTACGGAGCCAGTCTCTGACACTTGACATGCCAGCTCA 4052
Qy 1341 GluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLys 1360
Db 4053 GAGAGGGGGAAGGACACTGGAGGCTTGGAGGAAGCCCTTCTCTCCAGACCAAGCAAA 4112
Qy 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
Db 4113 GCAGGTGTGATAGAGTTCAAGTTCAAGGCTGACACACAGTCACTCAACAGCAGAA 4172
Qy 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
Db 4173 GCTGTGAAAGAGTTCATAGAAACGGTTGTGATTTCAAGACAGGTGAAGTCCAGAGTGT 4232
Qy 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrrThrLeu 1420
Db 4233 GTAGGTGCACACTTATACAGCTGAGAGTCTCTGCACGGGTGGCCACTGACTCTT 4292
Qy 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440
Db 4293 CAGCATGCAGAGGACACGGTACCCCTGGGCGCTGAGTCTCAGGCAGAAATCCATCCCAATC 4352
Qy 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyIleSerAla 1460
Db 4353 ATAGTAACCTCTGCTGAAAGACCCCTACCTCCTGACCTCAAGGAGAAATAAGCGCA 4412
Qy 1461 SerGlnArgGluArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
Db 4413 TCCAGAGAGCGCATCAGAGGAGAGGACAGCCAGATGCTGCTGTGATGCTGACGGC 4472
Qy 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
Db 4473 AAGGAGAGTACAGCAATCGAAAGATCCTCAAGCTGAACCTGAGATCCTGGAATTTGAG 4532
Qy 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
Db 4533 AGTAAGAGCAACAAAGATTGCTGAACTGCTCATTTCAGACAGCCGTTGACCAAGTTCGACGT 4592
Qy 1521 ThrGluThrAlaProGluThrHisAlaTytrAspSerGlnThrGlnValProAlaCysArg 1540
Db 4593 ACAGAAACACGCCCCGAAACTCATGCTTATGATTTCAGACACCCAGGTTCTGTGATGAGG 4652
Qy 1541 LeuAspSerArgGluProAsnArgCysTrrThrLysMetLysAspAlaLysMetLysHis 1560
Db 4653 CTTGACAGCAGGAGGCCAACACAGATGCTGGACAAATGAAGATGCCAAGATGAACAC 4712
Qy 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrrPalaGln 1580
Db 4713 CCAGTGGCGAGGCCAGAGAGGACTTGAAGTCTGACCTGACCTTCTGGAGGCGATGGCTCAG 4772
Qy 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 4773 CTTGGAAATGCTTGGCGCGCTTGCAGTTGAAAGCGCGCGGTGTCAAAG 4820

```

RESULT 3

US-08-665-401-3

; Sequence 3, Application US/08665401

; GENERAL INFORMATION:

; APPLICANT: Gelman, Irwin H.

; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE

```

; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,401
; FILING DATE: 18-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/34008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6160 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-665-401-3
;
; Alignment Scores:
; Pred. No.: 0 Length: 6160
; Score: 8073.00 Matches: 1596
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 100.00% Indels: 0
; DB: 10 Gaps: 0

```

US-09-902-432-4 (1-1596) x US-08-665-401-3 (1-6160)

```

Qy 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
Db 33 ATGGGCGCAGGCGAGTTCACCGGAGCAGCGAGCCCGGAGCGAGCGCGGAGCGACG 92
Qy 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
Db 93 CCGAGCGAGCTGGTGTCTCAGTGGCCATGGCCCGCAGCTGAAGCTCGGAGCGAGCTGGA 152
Qy 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
Db 153 GACCCCGCGCAGCGGACCCCGCCACCAAGCTCCACAGAGAATGGCCAGCTGTCTTCT 212
Qy 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsnGlnGluGlyGln 80
Db 213 GTCACCGCGTACGTGAACAGAGAGATGCCATGTCCACAGAGGAAACCCAGGAGGGGCGAG 272
Qy 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
Db 273 GAGCAAGAGAGTCTGTGATGAGGATGTTGGACAGCAGAGTCAAGAGATGTGAGAGAAAA 332
Qy 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120

```

[illegible]

QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal 860
DB 2553 GAGAAGATGAAGCCCAAGGGGAATACAGAGCTGCCAGCTGCTGGGGCTGTGTACGTG 2612
QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
DB 2613 TCCGAGAGCTCAGTAAGACTCTGTGTCACACTGTGAGTGTCCGAGTCATTGTATGGACC 2672
QY 881 ArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrGlu 900
DB 2673 AGGCGAGTCACAGTGTGGAAGAGCGGTCTCTTCGTGGATATCCCGTCCGTAACAGAA 2732
QY 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
DB 2733 CCTCTTGAACACACAGCGGAGAGCCATGCCACTCTTGAAGAGGTCACTGAAAAAGAC 2792
QY 921 IleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
DB 2793 ATCATTCAGAGAGAACTCTCTGTGCTCACCCAGACGTTACAGAGGGTAAAGATGCCCAT 2852
QY 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
DB 2853 GACGACATGCTCACCACTGAAGTGGATTTCACTCTCAGAACTGTGCAGCCACAGAGACC 2912
QY 961 SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAsp 980
DB 2913 TCAGAGGCTCTCCGTACTGAAGAAGTTTACCAGAACATCTGGGGGCCGAAGACACACAGAC 2972
QY 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
DB 2973 ATGGTGTCCGAGTTTCCCAGCTGACTGACTTCCCAGACACCACAGAGGAAGCCACCCCA 3032
QY 1001 ValGlnGluValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAla 1020
DB 3033 GTTCAGGAGTAGAGAGTGTGTGTAGATACAGAAAGAGGAGCGCCAGAGCC 3092
QY 1021 IleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThr 1040
DB 3093 ATCCTTCAAGCCGTTCAGACAAGGTGAAGAGAGTCCCAAGTGCCTGCAACCCAGACT 3152
QY 1041 ValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGlu 1060
DB 3153 GTCAGAGAACCGGGTCAAAAGACACTGGAGAAGTTGAGGAGGTAGAGGAGACTCCGAA 3212
QY 1061 ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
DB 3213 GTCTGGCTTCGGAGAAAGAGAGGAGCTTATGCCGAAAGACCCCGTCAGGAAGCTGGA 3272
QY 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
DB 3273 GCTGAGCATCTTGCACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTT 3332
QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
DB 3333 CCTGAAGTACCGCAGATGAGACCATGTCCGACGTCGAGGTTATCAAGCTCCAGCAG 3392
QY 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
DB 3393 CTGATGGAACAGCGCGTGGCCCTCGAGTCATCCGAAACCTTGACAGACAGTGAGACAAAT 3452
QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
DB 3453 GGAAGCAGCTCCCTTAGCAGATTTCAGACACTGCAGCTGCAGATGGGACACAGCAAGATGAACCAT 3512
QY 1161 AspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
DB 3513 GACAGCCAGACAGTAGTAAGCCACTGCAGCTGTTCAGCAGTTCACAGGTTCAGAGAAGAG 3572
QY 1181 AlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGlu 1200
DB 3573 GCGGCTACTGCTCAGAAAGAGGAGCCCTTCGACACTACTCTAATATGTTCCAGCCCAAGAA 3632
QY 1201 GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnLeuThrAla 1220

DB 3633 GAACATGGGGAAAGAACCCAGGAAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGCT 3692
QY 1221 AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeu 1240
DB 3693 GCAGCCGTGCCCGTCTCGCAAGACTGAGGTGGTCAAGAGGTGAGGTGTGACTGGTTG 3752
QY 1241 AspGlyGluLysValLysGluGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
DB 3753 GATGGAGAAAAAGTCAAGAAGAACAGGAGGTGTTGTACACTCTCGACCCCAACAGTCAA 3812
QY 1261 LysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
DB 3813 AAGGCTGCTGATGTGACATATGACAGTGAAGTGAATGGAGTGGCGGTGTCTAGGAAAAAG 3872
QY 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluGluMetGluThrAspValGlu 1300
DB 3873 GAGAGTACTGAGTGCAGAGTCTTAGCCTTGGAGAGGGAGATGGAACCTGACGTTGAA 3932
QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAla 1320
DB 3933 AAGCAAAAAGGGAGACAAAGCCAGAGCAAGTGAGTGAAGAGGTGAGCAGGAAAAACAGCC 3992
QY 1321 AlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSer 1340
DB 3993 GCTCCTGAGCATGAAGGAAACCTACGGGAAGCCAGTCTTGACACTTGACATGCCAGCTCA 4052
QY 1341 GluArgGlyLysAlaLeuGluGlySerLeuGlyGlySerProSerLeuProAspGlnAspLys 1360
DB 4053 GAGAGGGGAAGGACACTGGGAAGCCTTGGAGGAAGCCCTTCTCCAGACCAAGACAAA 4112
QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
DB 4113 GCAGGTTGCATAGAGTTCAGATTCAAGCTTCAAGCTGGACACAACTGACTCAAAACAGCAGAA 4172
QY 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
DB 4173 GCTGTGCAAAAAGGTCATAGAAAACGGTGTGATTTCAAGACAGAGTGAAGTCCAGAGTGT 4232
QY 1401 ValGlyValHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeu 1420
DB 4233 GTAGGTGACACACTATTACAGCTGAGAAGTCTCTGCAACCGGTGGCCACTGACTCTT 4292
QY 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440
DB 4293 CAGCATGCAGAGGACACACGCTACCCCTGGGGCCTGAGTCTCAGGACAGATCCATCCCAATC 4352
QY 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAla 1460
DB 4353 ATAGTAACTCCTGCTCCTGAAAGCACCTTACATCTCTGACCTACAGGAGAAAAATAAGCGCA 4412
QY 1461 SerGlnArgGluArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
DB 4413 TCCAGAGAGAGCGGATCAGAGGAAGAGGACAAAGCCAGATGCTGGTCTCTGATGTCGCGGC 4472
QY 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
DB 4473 AAGGAGGTACAGCAATCGAAAAAGTCTCAAGGCTGAACCTGAGATCTTGGAACTTGAG 4532
QY 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
DB 4533 AGTAAGAGCAACAAGATTGTGCTGAACGTCATTTCAGACAGCGCTTGACCACTTCGCACGT 4592
QY 1521 ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArg 1540
DB 4593 ACAGAAACAGCCCCGAAAACTCATGCTTATGATTTCACAGACCGGCTTGACCACTTCAGCAG 4652
QY 1541 LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
DB 4653 CTTGACAGCAGGAGGCCCAACAGATGCTGGACAAAATGAAGATGCCAAGATGAAACAC 4712
QY 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGln 1580

Db 4773 CCAAGTCCGACGACGAGAGAGACTTGCAAGCTCTGACCGCTTCTGAGGACATGGGCTCAG 4772
Qy 1581 ProArqlyscysleuProArqleuInleuYsaIaProValSerlys 1596
Db 4773 CCTCGGAATGCTTGCCGCGCTTGAGTTGAAAGCGCGGTGTCAAG 4820
RESULT 4
US-09-902-432-3
Sequence 3, Application US/09902432
GENERAL INFORMATION:
APPLICANT: Irwin H. Gelman
APPLICANT: Susan G. Jaken
TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
FILE REFERENCE: A30558-A-FWC-A 070156, 0597
CURRENT FILING DATE: 2002-04-08
PRIOR FILING DATE: 1997-11-25
PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: 08/665,401
PRIOR FILING DATE: 1996-06-18
PRIOR APPLICATION NUMBER: 08/635,121
PRIOR FILING DATE: 1996-04-19
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 6160
TYPE: DNA
ORGANISM: Rattus norvegicus
US-09-902-432-3
Alignment Scores:
Pred. No.: 0 Length: 6160
Score: 8073.00 Matches: 1596
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB:
US-09-902-432-4 (1-1596) x US-09-902-432-3 (1-6160)
Qy 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
Db 33 ATGGGCGCGAGCATTTCCACCGACGACGCGGAGCGCGCGGAGCGACAG 92
Qy 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
Db 93 CCGAGGAGACTGGTGTCTCAGTGGCCATGGGCCCGCAGCTGAAACCTTCGGGAGCAGCTGGA 152
Qy 41 AspProAlaAspAlaAspProAlaThrIlySleuProGluIlySasnGlyGluInleuSerSer 60
Db 153 GACCCCGCGACGCGGAGCCCGCCACCAAGCTCCACAGAAAGATGCGCAGCTGTCTTCT 212
Qy 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluIlySasnGlnGluGlyIn 80
Db 213 GTCAACGGCGTACTCAACAGAGATGTCTCATGTCCACAGAGAAAACAGAGGGGCGAG 272
Qy 81 GlnGluGlnValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluIly 100
Db 273 GAGGAAGAAAGTCTTATGATGATGTGGACAGCAGAGATCAGAAAGTGTGAGAGAAAA 332
Qy 101 AspArgValAlaGluIlyMetAlaAlaAsnSerThrAlaValGluAspIlyThrIlySasnGly 120
Db 333 GACCGAGTTGAAGAAATGGCGGCCCACTCCACAGCTGTGTAAGATATCAACAAAGGATGGG 392
Qy 121 GlnGluGlnThrSerGluIlyLeuGlnGlnIlyProAlaSerGluSasnValGluIly 140
Db 393 CAGGAGAGACATCAAGAAATATTGAACAGATCCCTGCTTCAAGAAACATGTGGAGAA 452
Qy 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheIlySlyValPheIlySlyPhe 160
Db 453 ATGGTACAGCGCTCTGATGCCAGAGCTAATGATGTGGCTTCAAGAAAGATTTAAATTT 512
Qy 161 ValGlyPheIlySlyPheThrValIlySlyAspIlySasnGluIlySerAspThrValGlnLeu 180

Db 513 GTTGCTTTAAATTCACGCTGAAGAGATATAAATGAAGTACATCTCTCAACTA 572
Qy 181 LeuThrValIlySlyAspGluGlyGluIlyAlaGluAlaSerValGlyAlaGlyAspHis 200
Db 573 CTCACTGTCAAGAAAGATGAAGAGGCGAGAGGCTCTGTCTGGAGCTGAGAGACAC 632
Qy 201 GlnGluProSerValGluThrAlaValGlyIlySerIlySasnGlySerGluIlySly 220
Db 633 CAGAGAGCCAGGTGGAGACATGCGCTGGAGAGTCACATCAAGAAAGTGAAGCTGAAG 692
Qy 221 GlnSerThrGluIlySlnGluGlnGlyThrLeuIlySlnGlnGlnSerSerThrGluIlyPro 240
Db 693 CAATCCACAGAGAGACAGAAAGGACCCCTGAAGCAGAGACAGAGACAGCAAGAAATCCCC 752
Qy 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGlnGluAlaIlySasnGlyGlyGluIlySly 260
Db 753 CTTCAAGCCGAATCTGATCAAGCGGCTGAGAGAGAACCAAGATGAAGAGAGAAAGAAAA 812
Qy 261 GlnGluIlySlnProThrIlySerProGluSerProSerSerProValAsnSerGluThr 280
Db 813 CAAGAGAAAGAGCCCAAGTCCCAAGATCCCGAGACGCCCACTCAACAGTGAAGCA 872
Qy 281 ThrSerSerPheIlySlyPhePheThrHisGlyTrpAlaGlyTrpArgIlySlyThrSer 300
Db 873 ACATCTTCTTCAAGAAATCTTCACTCAAGCTTGGCGGCTGGCGGCAAGAACCAAGCAGC 932
Qy 301 PheIlySlySerIlySlnAspIlySlnGluThrAlaGluIlySasnGlyGluGlnGluAla 320
Db 933 TTCAAGAAATCAAAAGAGATGATCTGGAATCTGCCGAGAAAGAGAGAGCAAGAGCA 992
Qy 321 GlnIlySlyValAspGluGlnGluIlySlnGlySlyThrGluProAlaSerGluGlnGluPro 340
Db 993 GAAAAAGTGAAG 1052
Qy 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspIlyGlyIlySlyValGluIlyLeuPro 360
Db 1053 GCGAGAGACACAGACACAGCCAGCTTGTACAGACATCAGAGAGAGTGAAGTGTGCTTTG 1112
Qy 361 GluAspGlnValGlyAspIlySlnGluIlySasnSerGluIlySlySasnProLeuAlaThr 380
Db 1113 GAAGACAGAGTGTGATCTGGAGGATGTCTCAGAGAGAGAGTGTCTTGTGGCAAG 1172
Qy 381 GluValPheAspGluIlySlyMetGluAlaHisGlnGlnValAlaAlaGluValHisValSer 400
Db 1173 GAAGTGTGTGATGAAGAGATGAAGCCCAAGAGAGTGTGTGAGAGATCACTGAGC 1232
Qy 401 ThrValGluIlySlyThrGluGlnGlnGlyGlyGlyGlyGluAlaGluIlyGlyValVal 420
Db 1233 ACCGTGGAGAGACAG 1292
Qy 421 ValGluGlyThrGlyGluSerLeuProGluIlySasnAlaGluProGlnGluValPro 440
Db 1293 GTAGAAGGAACAGAGAAATCTTGGCCCTTGAGAAATCTGAGAGAGAGAGAGAGAG 1352
Qy 441 GlnGluAlaGluProAlaGluIlyMetIlySlySasnArgGluMetCysValSerGlyGly 460
Db 1353 CAGGAAGCTGAGCTGTGAGAGATGATGAAGACAGAGAGATGTGTCTCTGAGAGA 1412
Qy 461 AspHisSerThrGlnLeuThrAspLeuSerProGluGluIlySlyThrLeuProIlyHisPro 480
Db 1413 GACCACTCACTCAACAGACCTTAAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
Qy 481 GlyIlyValSerGluValGluMetLeuSerSerGlnGluIlySlyValGlnGlySer 500
Db 1473 GGCATTGTCAAGTGAAGTGAAGTGTGTCTCTCAAGAAAGATCAAGGTACAGGAGAT 1532
Qy 501 ProLeuIlySlyLeuPheSerSerSerGlyIlySlySlySlySerGlyIlySlySlnGly 520
Db 1533 CCTTGAAGAAATCTTCAAGTGTGATGAGCTTAAAGAGAGTCTTGGGAGAGAGAGAG 1592
Qy 521 GlyIlySasnGlyGlyGlyValIlySasnGluIlyProGluIlyGluIlyGlnHisIlySerThr 540


```

QY 1261 LysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
Db 3813 AAGGCTGCTGATGTGACATATGACAGTGAAGATGGAGTGGCGGGGTGACGAGAAAAG 3872
QY 1281 GtuserThrGluValGlnSerLeuSerLeuGluGluGluValMetGluThrAspValGlu 1300
Db 3873 GAGAGTACTGAAAGTGCAGAGCTTACCTCGAGAGAGAGAGATGGAACCTGACCTGGA 3932
QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAla 1320
Db 3933 AAGGAGAAAAGGAGACAAAGCCAGAGCAAGTGAAGAAAGTGAAGAGGAGGAAACAGCC 3992
QY 1321 AlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSer 1340
Db 3993 GCTCTGACATGAAAGAACTTACGGGAAAGCCAGTCTTACACTTGAACATGCCCCAGCTTA 4052
QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyGlyLysProSerLeuProAspGlnAspLys 1360
Db 4053 GAGAGGGGAGAGGACATGGGAAACCTTGAAGAGAGCCCTTCTCTCCAGACCAAGACAAA 4112
QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
Db 4113 GCAGGTTGATAGAGGTTCAGAGTCAAGCCTGAGACAAACAGTCACTCAAAACAGAGAA 4172
QY 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
Db 4173 GCTGTGAAAAGGTCAATAAACGGTTGATTTCCAGAGACAGGTGAAGATCCAGAGTGT 4232
QY 1401 ValGluAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyIleThrThrLeu 1420
Db 4233 GTAGGTGCACACTTATTCACAGCTGAGAAAGTCTCTGCAAGCGGTGGCCACTGACTCTT 4292
QY 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440
Db 4293 CAGCATGACAGAGACACGGTACCCCTGGGGCCCTGAGTCTCAGGCAAAATCCATCCCAATC 4352
QY 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnIleGlyIleSerAla 1460
Db 4353 ATAGTAACCTCGCTCTCTGAAACACCCCAACATCCGACCTTAACAGAGAAATACCGCA 4412
QY 1461 SerGlnArgGluAspSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
Db 4413 TCCCAAGAGAGAGCATCAAGAGAGAGAGACACAGATCTGCTGCTGATGCTGACGCGC 4472
QY 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
Db 4473 AAGGAGAGTACAGCAATCGAAAAAGTCTCAAGGCTGAACTGAGATCCGGAACCTTGAG 4532
QY 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
Db 4533 AGTAAAGAGCAACAAATGTGTGAAAGCTCATTCACAGACGCGTTGACCGATTCGACAGT 4592
QY 1521 ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArg 1540
Db 4593 ACAGAAACAGCCCCCAAACTCATGCTTATGATTCACAGACCCAGGTCTCTGATGACAG 4652
QY 1541 LeuAspSerArgGluProAsnArgCysTyrPThrLysMetLysAspAlaLysMetLysHis 1560
Db 4653 CTGACAGAGAGGAGCCCAACAGATCTGGAACAAAATGAAGATGCCAAGATGAAGAACAC 4712
QY 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluValAlaThrPAlaGln 1580
Db 4713 CCGATCCCGAGCCACAGAGAGGAGCTTGCAAGTCTTGAACCGTTGAGGAGATGGGCTCAG 4772
QY 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 4773 CCTCGAAATGCTTCCGCGCTTGCAAGTGAAGCCCGGATGTCAGAG 4820

```

RESULT 5
 PCT-US02-16173A-2063
 ; Sequence 2063. Application PC/TUS0216173A
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendrick, Donna

```

; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-MO
; CURRENT APPLICATION NUMBER: PCT/US02/16173A
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2063
; LENGTH: 5236
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_057103
PCT-US02-16173A-2063

Alignment Scores:
Pred. No.: 0 Length: 5236
Score: 8030.00 Matches: 1589
Percent Similarity: 99.75% Conservative: 3
Best Local Similarity: 99.56% Mismatches: 4
Query Match: 99.47% Indels: 0
DB: 1 Gaps: 0

US-09-902-432-4 (1-1596) x PCT-US02-16173A-2063 (1-5236)
QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
Db 33 ATGGCGGAGAGGAGTTCACAGCAGAGCGAGAGCCCGGAGCGGAGCGACAGCAG 92
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyValAlaGly 40
Db 93 CCGAGCGAGCTGTCTCTGATGCGCATGAGCCCGAGCTGAAAGCTCCGGAGACACTGGA 152
QY 41 AspProAlaAspAlaAspProAlaThrLysLeuProGluLysAsnGlyGlnLeuSerSer 60
Db 153 GAGCCCGCGAGCGGAGCCCGCCACCAAGCTCCACAGAAAGATGGCAGCTGTCTTCT 212
QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGluGln 80
Db 213 GTCAACGGGTGATGGAACAAGAGATGTCCATGTCCAAAGAGAAAACAGAGAGGCGAG 272
QY 81 GluGluGluValValAspGluAspValAlaGlyGlnArgGluSerGluAspValArgGluLys 100
Db 273 GAGGAAGAAGTGTGTGATGAGATGTGACACGCAAGATCGAAGATGTGAGAGAAAAA 332
QY 101 AspArgValAlaGluMetAlaAlaAsnSerThrAlaValAlaGluAspIleThrLysAspGly 120
Db 333 GACCGAGTTGAAGAAATGGCGCCCAACTCCACAGCTGTGTGAAGATATCACAAAGATGG 392

```

QY 121 GlnGluThrSerGluIlelleGluGlnIleProAlaSerGluAenAenValGluGlu 140
DB 393 CAGGAGAGACATCAGAAATAATTTGAACAGATCCCTCTTCAGAAACAATGTGGAAGA 452
QY 141 MetValGlnProAlaGluSerGlnAlaAenAspValGlyPheIlyeLysValPheLysPhe 160
DB 453 ATGGTACAGCGCTGAGTCCAGGCTTAATGATGTTGGCTTCAAGAAAGTATTAAATTT 512
QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
DB 513 GTTGGTTTAAATTCACGGTGAAGAGGATAAAAAATGAAAGTCAGATACTGTCCAACTA 572
QY 181 LeuThrValLysLysAspGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
DB 573 CTCACCTGTCAAGAAAGGATGAAGCGCAGAGGGCAGAACCTCTGTCCGAGCTGGAGACCAC 632
QY 201 GlnGluProSerValGluThrAlaValGlyLysSerAlaSerLysGluSerGluLeuLys 220
DB 633 CAGAGCCCAAGTGTGGAGACTGCGCTCGGAGAGTCAGCATCCAAAGAAATGAGCTGAAG 692
QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIlePro 240
DB 693 CAATCCACAGAGAAAGCAAGGACACCTCTGAAGCAAGAACAGAGCAGCACAGAAATCCCC 752
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLys 260
DB 753 CTTCAAGCCCAATCTGATCAAGCGCTGAGCAAGAACCCAAAGATGAAGAGAGAAATAA 812
QY 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAenSerGluThr 280
DB 813 CAAGAGAAAGAGGCCCAACCAAGTCCCCAAGATCCCGAGCAGCCAGCTCAACAGTGAGACA 872
QY 281 ThrSerSerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSer 300
DB 873 ACATCTCTCTTCAAGAAAGTTCTCACTACGGTTGGCCCGGCTGGCCCAAGAGACCAGC 932
QY 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
DB 933 TTCAAGAAATCAAAAGAGGATGATCTGAAACTGCCGAGAAAGAGAGAGAGAGGCA 992
QY 321 GluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGlnGluPro 340
DB 993 GAAAAAGTAGACGAGGAG 1052
QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeu 360
DB 1053 GCAGAGACACACACCCAGGCGAGGTTGTACGACACTACGAGAGGTTGGAGCTGCCCTTG 1112
QY 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThr 380
DB 1113 GAAGACCAGGTTGGTACCTGGAGGCATCTGCAGAGAGAGAGTGTGCTCTTTGGCAAG 1172
QY 381 GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
DB 1173 GAAGTGTGTTGATGAGAGATGGAAGGCCACCAAGAACTGTTGCAGAGGTTCCACGTGAGC 1232
QY 401 ThrValGluLysThrGluGluGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyValVal 420
DB 1233 ACCGTGAGAGAGACAGAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1292
QY 421 ValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValPro 440
DB 1293 GTAGAGAGACAGAGAGATCTTTGCCCTCGAGAACTGCTGAGCCCAAGAGGTTCCCC 1352
QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
DB 1353 CAGGAGCTGAGCGTGTGAGGAGCTGATGAAGAGCAGAGAGATGTGTGTCTCTGGAGGA 1412
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
DB 1413 GACCACACTCACTGACAGACCTTAAGTCTCTGAAGAGAGAGAGCTGCCCAACACCCAGAA 1472

QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500
DB 1473 GGCATTGTTCAGTTCAGTGGAGATGCTGCTCTCAGGAAAGAAATCAAGGTACAGGGAAT 1532
QY 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
DB 1533 CCCTTGAGAGAACTCTTCAGTACTCAGGCTTAAAGAACTGTCTGGGAAGAACAGAG 1592
QY 521 GlyLysArgGlyGlyGlyGlyAspGluGluProGlyGluTyrGlnHisIleHisThrGlu 540
DB 1593 GGGAAACAGGAGGAGTGGCGGAGACGAGAGAGCTGCGGTCTGCCCGAGGAGCCT 1652
QY 541 SerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluPro 560
DB 1653 TCCCCAGAGAGTCTGATGAGCAGAGAGGAGAGAGCTCTGCGTCTGCCCGAGGAGCCT 1712
QY 561 GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
DB 1713 GAGGAGACCACGTGTCTGGAGAAAGGCGCTGGAGAACCCCGAGATGGGAAAGCTGAG 1772
QY 581 GluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAlaSerPhe 600
DB 1773 GAAGGAACCTACTTCGATGGAGAGAGAGAGAGAGGATCACTCCCTGGGCATCTCTTC 1832
QY 601 LysLysMetValThrProLysLysArgValArgArgProSerSerGluSerAspLysGluGlu 620
DB 1833 AAAAAGATGTTGACACCCCAAGAAACGGGTCCGAAGACCTTCTGAGAGTGACAGAGGAA 1892
QY 621 GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
DB 1893 GAGCTGGAGAAAGTCAAGAGCGCCACTTGTCTCTCACTGTAGTACAGCTGTTCAGAAATG 1952
QY 641 GlnAspGluValLysThrValGlyGluGlnLysProGluProLysArgArgVal 660
DB 1953 CAAGATGAAGTCAAACTGTTGGTGAAGAACAAAGCCAGAGAGAACCAAGACGTAGGGTG 2012
QY 661 AspThrSerValSerTrpGluAlaLeuLysCysValGlySerSerLysLysArgAlaArg 680
DB 2013 GATACTTCAGTGTCTTGGGAAGCACTGATTTGTGTCGATCATCCAAAGAGAGAGCAAG 2072
QY 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
DB 2073 AAGGCATCTCTTCAGATGATGAAGAGGAGGCGCAAGGACACTGGGAGGGGACAGTCAAGA 2132
QY 701 AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
DB 2133 GCAGAGAGGCGCAGCAAGAACCAAGAGCCGGAACAGAGCGTGTCTCTGCCAGCACCCAG 2192
QY 721 GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly 740
DB 2193 GAGCAGGACCAAGCGCAAGAAAGTTCTCCACCGAGCCAGCGGGAAGCCTTCCGAAGGG 2252
QY 741 GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
DB 2253 GAAGTGTCTCCACTTGGGAGTCATTAAAGATTTAGTCACCTCCAAAGAAAAAATCCAAG 2312
QY 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
DB 2313 TCAAAACCTGGAAGAGAAAGCCGAAGACTCTAGTGTAGACAGTGTGTCCACTGAGATCGAA 2372
QY 781 ProSerArgGluGluSerTrpValSerLysLysPheIleProGlyArgArgLysLys 800
DB 2373 CCGAGTAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGGCGGAGAGAA 2432
QY 801 ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsn 820
DB 2433 AGGGCAGATGGGAAGCAAGAACCAAGCTGTGGAAGACTCAGGGCCAGTGGAGATAAAT 2492
QY 821 GluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaValGluArg 840
DB 2493 GAGGAGACCCCTGATGTCCAGCCGTCGTCCTCTGAGTATGATGTCAGTGGAGAG 2552
QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrVal 860

Db 2553 GAGAAAGATGGAAGCCAGGAGGAATGGGAGCTCCCTCCAGCTGCTGGGGCTGTATCGTG 2612
 QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
 Db 2613 TCCGAGGAGACTCATGTAAGACTGTGTCCACACTGTAGTGTCCAGTCACTTGTGGAGACC 2672
 QY 881 ArgAlaValThrSerValGluGluGluGluSerProSerThrIleSerAlaSerValThrGlu 900
 Db 2673 AAGGAGCTACACAGTGTCCAAAGGCGGTCTCTTGTGATATCCCTTCCGTAACAGAA 2732
 QY 901 ProLeuGluHisThrAlaGluGluAlaMetProProValGluGluValThrGluLysAsp 920
 Db 2733 CCTCTTGAACACACAGCGGAGAGAGCATGCCCTCTTGAAGAGGTCACTGAAAAAGAC 2792
 QY 921 IleIleAlaGluGluThrProValLeuThrGluThrLeuProGluGluLysAspAlaHis 940
 Db 2793 ATCATTTGCAAGAAACTCTGTGTCTCACCCAGAGCTTACAGAGGGTTAAAGATGCCCAT 2852
 QY 941 AspAspMetValThrSerGluValAlaAspPheThrSerGluAlaValThrAlaThrGluThr 960
 Db 2853 GACGACATGTCACACAGTGAAGTGATTTCACTCAGAAAGCTGTGACAGCCACAGAGACC 2912
 QY 961 SerGluAlaLeuArgThrGluGluValThrGluAlaSerGluValAlaGluGluThrThrAsp 980
 Db 2913 TCAGAGGCTCTCGTACTGAAGAGTTACCGAAGCATCGGGGCGCCAGAGACCAAGAC 2972
 QY 981 MetValSerAlaValSerGluLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
 Db 2973 ATGGTGTCCGCAATTTCCACAGTGACTGACTCCCGACAGACCAAGAGGAGCCACCCCA 3032
 QY 1001 ValGluGluValGluSerGlyValLeuAspThrGluGluGluGluThrGluThrGluAla 1020
 Db 3033 GTTCAGAGAGTGAAGAGTGTGTGTAGATACAGAAAGAGAGAGGCGCCAGAGCGCGCC 3092
 QY 1021 IleLeuGluAlaValAlaAspLysValLysGluGluSerGluValProAlaThrGluThr 1040
 Db 3093 ATCTCTCAAGCCCTTCCAGACAGGTGAAGAGAGAGTCCAGGTGCTGCAACCCAGACT 3152
 QY 1041 ValGluArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGlu 1060
 Db 3153 GTGCAGAGAAAGCGGGTCAAAAGCATGGAAGAGTTGAGAGAGTTAGAGAGGAGCTCCGAA 3212
 QY 1061 ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGluGluAlaGly 1080
 Db 3213 GTGCTGGCTTCGAGAAAGAGAGAGCGTTATGCCAAAGAGACCCGTGACGAGAGCTGGA 3272
 QY 1081 AlaGluHisLeuAlaGluGluSerGluThrGluGluAlaThrProGluSerLeuGluVal 1100
 Db 3273 GCTGAGCATCTTGACAGAGGCTCTGAGACTGACAGAGCTACTCCAGAGACCTTGAAGTT 3332
 QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGluGluValIleLysLeuGluGlu 1120
 Db 3333 CCTGAAGTACCGGCAATGTAGACATGTGTGCCAGCTGACAGGTTATCAAGCTCCAGCAG 3392
 QY 1121 LeuMetGluGluAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
 Db 3393 CTGATGGAACAGGCGCTGGCGCTCTGAGTCAATCCGAAACCTTGAACAGACGTGAGACAAAT 3452
 QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGluThrGluGluGluAspGluThrIle 1160
 Db 3453 GGAAGAGACTCCCTTACACAGTTACAGACACTGACGATGGGACACAGCAAGATGAAACCATT 3512
 QY 1161 AspSerGluAspSerLysAlaThrAlaAlaValArgGluSerGluValThrGluGluGlu 1180
 Db 3513 GACAGCCAGGAGACGTAAAGCACTGCACAGTGTCAAGGACGTCAACAGTCAAGAAAGAG 3572
 QY 1181 AlaAlaThrAlaGluLysGluGluProSerThrLeuProAsnAsnValProAlaGluGlu 1200
 Db 3573 GCGGCTATCTGCTCAGAAAGAGAGCCTTGCACACTATATATGTTCAGGCCCGAGAA 3632
 QY 1201 GluHisGlyGluGluProGluArgAspValLeuGluProThrGluGluGluLeuThrAla 1220

Db 3633 GAACATGGGGAAGAACCAAGAGATGTTCTTGAACCTTACACAGCAAGCTTACTGCT 3692
 QY 1221 AlaAlaValProValLeuAlaLysThrGluValGluGluGluGluGluValAspThrLeu 1240
 Db 3693 GCAGCCGTGCCCTTCTGCGCAAAAGACTGAGTGGGTCAAGAGGGGTGAGCTGACTGTG 3752
 QY 1241 AspGlyGluLysValLysGluGluGluGluValPheValHisSerGlyProAsnSerGlu 1260
 Db 3753 GATGAGAAAAAGTCAAAAGAAAGACAGAGGTGTTTGTACCTCTGGACCCCAAGCTGAA 3812
 QY 1261 LysAlaAlaAspValThrTyraAspSerGluValMetGlyValAlaGlyCysGluGluLys 1280
 Db 3813 AAGCTGCTGATGTACATATGACAGTGAAGTGAAGTGGAGTGGCGGGGTTCAGGAAAG 3872
 QY 1281 GluSerThrGluValGluSerLeuSerLeuGluGluGluGluMetGluThrAspValGlu 1300
 Db 3873 GAGACTACTGAAGTCCAGAGTCTTACCTGGGAGGAGAGATGGAACCTGACCTTGA 3932
 QY 1301 LysGluLysArgGluThrLysProGluGluValSerGluGluGluGluGluThrAla 1320
 Db 3933 AAGGAGAAAAAGGAGACAAAGCCAGAGCAAGTGAAGAGTGAAGAGTGAACAGAAACAGCC 3992
 QY 1321 AlaProGluHisGluGlyThrTyrgLysProValLeuThrLeuAspMetProSerSer 1340
 Db 3993 GCTCTGAGCATGAAGGAACCTACGGGAAGCCAGTCTGACACTTGACATGCCAGCTCA 4052
 QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGluAspLys 1360
 Db 4053 GAGAGGGGGAAGGCACTGGGAAAGCTTGGAGAGAGCCCTTCTCTCCAGACCAACAGAAA 4112
 QY 1361 AlaGlyCysIleGluValGluValGluSerLeuAspThrThrValThrGluThrAlaGlu 1380
 Db 4113 GCAGGTGTGATAGAGTTCAGATTCAAGCTTGACACAAAGTACTCAAAACACAGAA 4172
 QY 1381 AlaValGluLysValIleGluThrValIleSerGluThrGlyGluSerProGluLys 1400
 Db 4173 GCTGTGAAGAAAGCATATGAAGAAACGTTGTGATTTCAAGACAGAGTGAAAGTCAAGAGT 4232
 QY 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisIleThrLeu 1420
 Db 4233 GTAGGTGCACTTATTTCACAGCTGAGAAAGTCTCTGCAACGGGTGCACATGACTCTT 4292
 QY 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGluAlaGluSerIleProIle 1440
 Db 4293 CAGCATGCAAGAGAACCGGTACCTCGGGCGCTGTGATCTCAGGCAAGATCCATCCAAATC 4352
 QY 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGluGlyLysIleSerAla 1460
 Db 4353 ATAGTAACTCCCTGCTCTGAAAGCAACCTTCACTGACCTTACAGAGGAATAAGCGCA 4412
 QY 1461 SerGluArgGluArgSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
 Db 4413 TCCCAAGAGAGCGATTCAGAGGAAAGAGCAAGCCAGATGTGCTGTCTGATCTGACGGC 4472
 QY 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluLysLeuGluLeuGlu 1500
 Db 4473 AAGGAGAGTACAGCAATGAAAAAGTCTCAAGGCTGAAACCTGAAGATCTTGAACTTGAG 4532
 QY 1501 SerLysSerAsnLysIleValLeuAsnValIleGluThrAlaValAspGluPheAlaArg 1520
 Db 4533 AGTAAGAGCAAAAGATTTGTCTGAACGTCAATTCAGACAGCCGTGTGACCAAGTTCACAGT 4592
 QY 1521 ThrGluThrAlaProGluThrHisAlaTyraAspSerGluThrGluValProAlaCysArg 1540
 Db 4593 ACAGAAACAGCCCCGAAACCTCATGCTTATGATTCACAGACCAGAGTTCTTGCAATGCGG 4652
 QY 1541 LeuAspSerArgGluProAsnArgCysIleThrLysMetLysAspAlaLysMetLysHis 1560
 Db 4653 GCTGACAGAGGAGGCCAAACAGATGTGGACAAAATGAAGTTGCCAAGTGAAGAACAC 4712
 QY 1561 ProValProGluProArgGluAspLeuGluValLeuThrValLeuGluValAlaThrAlaGlu 1580
 Db 4713 CCAAGTCCGCCACCCAGAGAGGACTTGCAGAGTCTGACCGTTCTTGAGGCAATGGGCTCAG 4772

QY 1581 ProArgLysCysLeuProArgLeuGlnLeuAlaProValSerLys 1596
|
Db 4773 CTCGGAAATGCTTGGCGCGCTTGCAGTTGAAAGCGCGGTGCAAG 4820

RESULT 6

US-10-191-803-279
; Sequence 279, Application US/10191803
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 279
; LENGTH: 5236
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_057103
US-10-191-803-279

Alignment Scores:

Pred. No.: 0 Length: 5236
Score: 8030.00 Matches: 1589
Percent Similarity: 99.75% Conservative: 3
Best Local Similarity: 99.56% Mismatches: 4
Query Match: 99.47% Indels: 0
DB: 41 Gaps: 0

US-09-902-432-4 (1-1596) x US-10-191-803-279 (1-5236)

QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
|
Db 33 ATGGGCGCAGGCAGTTCACCGAGCAGCGAGCGGCCCGCGCGGAGCGACAG 92
|
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
|
Db 93 CCAGCGAGCTGGTGCTCAGTGGCCATGGGCCCGCCGACCTGAGCCCTCGGAGCAGCTGGA 152
|
QY 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
|
Db 153 GACCCCGCCGACGCGGACCCCGCCACCAAGCTCCACAGAAAGATGSCCAGCTGTCTTCT 212
|
QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsnGlnGlyGln 80
|
Db 213 GTCAACGCGGTAGCTGAACAGGAGATGTCCATGTCCAAGAGGAAACCAGGAGGGCGAG 272
|
QY 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
|
Db 273 GAGGAGAGTCTGTTGATGAGGATGTTGGACAGGAGATGAGAGATGTGAGAGAAAA 332
|
QY 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
|
Db 333 GACCGAGTTGAAGAAATGGCGGCCAATCCACAGCTGTTGAGAGTATCACAAGGATGGG 392
|
QY 121 GlnGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnValGluGlu 140
|

Db 393 CAGGAGGAGACATCAGAAAATAATTGAACAGATCCCTGCTTCAAAAAACAATGTGGAGAA 452
|
QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
|
Db 453 ATGGTACAGCCTGCTGAGTCCCAGGCTAATGATGTTGGCTTCAAGAAAGTATTATAATTT 512
|
QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
|
Db 513 GTTGGTTTTAAATTCACGGTGAAGAGGATAAAAAATGAAAGATCAGATACTGTCCAACTA 572
|
QY 181 LeuThrValLysLysAspGluGlyGluAlaGluAlaSerValGlyAlaGlyAspHis 200
|
Db 573 CTCACCTGTCAAGAGGATGAAGCGGAGGGCGAAGACCTCTGTGCGAGCTGGAGACCAC 632
|
QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
|
Db 633 CAGGAGCCAGTGTGAGAGTCCGCTCGAGAGTCAAGCATCCAAAGAAAGTAGCTGAAG 692
|
QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIlePro 240
|
Db 693 CAATCCACAGAGAGCAAGAACGCCACCTCAAGCAAGAAACAGACGACGACAGAAATCCCC 752
|
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluLys 260
|
Db 753 CTTCAAGCCGATCTGATCAAGCGCTGAGGAAGAACCAAGATGAAGAGAAAGAAAA 812
|
QY 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThr 280
|
Db 813 CAAGAGAAAGAGCCCAACCAAGTCCCCCAGAAATCCCCGAGCAGCCCACTCAACAGTGACA 872
|
QY 281 ThrSerSerPheLysLysPheThrHisGlyTrpAlaGlyTrpArgLysLysThrSer 300
|
Db 873 ACATCTTCTTCAAGAAAGTTCTTCACTACGGTGGCGCGCTGGCGCAAGAGACGACG 932
|
QY 301 PheLysLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
|
Db 933 TTCAGAAATCAAAAGAGGATGATCTGGAAACTGCCGAGAGAGAAAGGAGAGAGGCA 992
|
QY 321 GluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGluGlnGluPro 340
|
Db 993 GAAAAAGTAGACGAG 1052
|
QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeu 360
|
Db 1053 GCAGAAAGACACAGACCAGCCAGGTTGTCTACGACACTACGAGAAAGGTGGAGCTGCTTTG 1112
|
QY 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThr 380
|
Db 1113 GAAGACCAGGTTGGTGCACCTGGAGGAGCATCGTCAGAGGAGAGTGTCTCTTTGGCAACG 1172
|
QY 381 GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
|
Db 1173 GAAGTGTGTGATGAGAGATGGNAGCCCAACCAAGAAAGTTGTCAGAGGTCCACGTGAGC 1232
|
QY 401 ThrValGluLysThrGluGluGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyValVal 420
|
Db 1233 ACCGTGGAGAGACAGAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1292
|
QY 421 ValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGluValPro 440
|
Db 1293 GTAGAGGAGAACAGGAGAAATCCTTTGCCCTCTGAGAAACTGGCTGAGCCCGCAGGAGGTC 1352
|
QY 441 GlnGluAlaGluProAlaGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
|
Db 1353 CAGGAAGCTGAGCCTCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAGGA 1412
|
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
|
Db 1413 GACCACACTCACTGACAGACCTAAGTCTCTGAAGAGAGAGCGCTGCCCAACACCCAGAA 1472
|
QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500
|
Db 1473 GGCATTGTGAGTGGAGATGCTGTCTCTCTCAGGAAAGAAATCAAGGTACAGGGAAGT 1532
|

QY 501 ProLeuLysLeuPheSerSerSerGlyLeuLysLeuSerGlyLysLysGlnLys 520
Db 1533 CCCTTGAAGAACTCTTCTGATGCTCAAGGCTTAAAGAACTGTCTGGAGAAAGCAAG 1592
QY 521 GlyLysArgGlyGlyGlyLysAspGlnProGlyGlyLysGlnHisLeuSerGln 540
Db 1593 GGGAAACGAGAGGTGGGAGAGCAAGAGCTCGAAGATACCAACACTTACACCGAA 1652
QY 541 SerProGlnSerAlaAspGlnLysGlyGlnSerSerAlaSerSerProGlnLysPro 560
Db 1653 TCCCAAGAGAGTGGATGATAGCAAGAGGAGAGAGCTCGCTCGTCCCCCGAGAGCCT 1712
QY 561 GlnGlnThrThrCysLeuGlnLysGlyProLeuGlnLysAlaProGlnAspGlyLysAlaGln 580
Db 1713 GAGGAGACCACTGTCTGAGAAAGGCGCTGGAAGCACCCAGAGATGGGAGCTGAG 1772
QY 581 GlnGlyThrThrSerAspGlyGlyLysLysArgGlnGlyIleThrProThrAlaSerPhe 600
Db 1773 GAAGGAATCTCTCCGATGGAGAGAAAGAGAGAGGATCACTCTCGGCACTCTTC 1832
QY 601 LysLysMetValThrProLysLysArgValArgArgProSerGlnSerAspLysGlnLys 620
Db 1833 AAAAAGATGTGACACCCCAAGAAACGGGTCCGAAGACTTCTGAGAGTGACAAAGAGGAA 1892
QY 621 GlnLeuGlnLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGlnMet 640
Db 1893 GAGCTGGAGAGAGGTCAAGAGCCCACTTGTCTCCACTGATGACACACTGTCTAGAAATG 1952
QY 641 GlnAspGlnValLysThrThrValGlyGlnGlnLysProGlnLysProLysArgArgVal 660
Db 1953 CAAGATGAAGTAAACGTGTGTGAGAAACAAAGCCAGAGAACCAAGGCTAGGGTG 2012
QY 661 AspThrSerValSerThrGlnLysLeuLysCysValGlySerSerLysLysArgAlaArg 680
Db 2013 GATCTTCACTGCTCTTGGAGACACACTGTTGTGTGGATCATCCAAAGAGAGCAAG 2072
QY 681 LysAlaSerSerSerAspAspGlnGlyGlyProArgThrLeuGlyGlyLysSerHisArg 700
Db 2073 AAGGATCTCTTCTCAGATGATGAAGAGGGCCAAAGGACCTGGAGGGGACGTCCACA 2132
QY 701 AlaGlnGlnLysSerLysAspLysGlnAlaGlyThrAspAlaValProAlaSerThrGln 720
Db 2133 GAGAGAGAGGCCAGCAAGACAAAGAGCCGGAACAGAGCTGTCTCTCCAGCACCCAG 2192
QY 721 GlnGlnAspGlnAlaGlnLysSerSerSerProGlnProAlaGlySerProSerGlnGly 740
Db 2193 GAGCAGAGACCAAGCCGCAAGAGTTCTCACCCGAGCCAGCGGAAAGCCCTTCCGAAGG 2252
QY 741 GlnGlyValSerThrThrGlnSerPheLysArgLeuValThrProArgLysLysSerLys 760
Db 2253 GAAGGTGTCTCCACTTGGAGATCATTTAAAGATTAGTCACTCCAAAGAAAAATCCAA 2312
QY 761 SerLysLeuGlnGlnLysAlaGlnLysSerSerValGlnLysLeuSerThrGlnLysGln 780
Db 2313 TCAAAACCTGGAGAGAAACCCAGAAAGCTAGTGTAAGACAGTGTGCCACTGAGATCGAA 2372
QY 781 ProSerArgGlnGlnSerThrValSerLysLysPheIleProGlyArgArgLysLys 800
Db 2373 CCGATGAGAAAGATCTTGGGTTTCCATTAAAGAAATTCATCCCGGAGCGCGAGAAAGAA 2432
QY 801 ArgAlaAspGlyLysGlnGlnGlnAlaThrValGlnAspSerGlyProValGlnLysLeu 820
Db 2433 AAGGCGAGATGGGAGCAAGAACAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATTAAT 2492
QY 821 GlnAspAspProAsnValProAlaValValProLeuSerGlnLysAsnAlaValGlnArg 840
Db 2493 GAGGAGCAGACCTGATGTCTCCAGCCGCTGCTGTCTGAGATGATGATCAGTGGAGAG 2552
QY 841 GlnLysMetGlnAlaGlnGlnLysAnThrGlnLysProGlnLysLeuGlnLysAlaValLysVal 860
Db 2553 GAGAAAGATGAAGCCCGAGGGAATGCGAGCTGCCCACTGCTGGGGCTGTGTACGTG 2612

QY 861 SerGlnLysLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
Db 2613 TCCGAGAGCTCAGTAAGACTCTGATCCACACTGTGATGTGCGAGTATGATGAGGACC 2672
QY 881 ArgAlaValThrSerValGlnGlnLysArgSerProSerThrIleSerAlaSerValThrGln 900
Db 2673 AAGGCACTACCAAGTGTGGAAGAGGGGTCTCTCTGTGTGATATCCGCTTCCGTTAACAA 2732
QY 901 ProLeuGlnHisThrAlaGlyGlnLysMetProProValGlnGlnValThrGlnLysAsp 920
Db 2733 CCTCTTGAACACACAGCCGGAGAACCTGACCTGCTGTGAAGAGTCACTGAAGAAAC 2792
QY 921 IleIleAlaGlnGlnThrProValLeuThrGlnThrLeuProGlnGlyLysAspAlaHis 940
Db 2793 ATCATTTGAGAAAGAACTCTGTGTCTACCCAGAGGTATCCAGAGGGTAAAGATGCCAT 2852
QY 941 AspAspMetValThrSerGlnValAspPheThrSerGlnAlaValThrAlaThrGlnThr 960
Db 2853 GACGACATGTCTCACCGAGAGATGGAATTTCACTCAGAAAGCTGTGACAGCCACAGAGACC 2912
QY 961 SerGlnAlaLeuArgThrGlnGlnValThrGlnLysArgValAlaGlnGlnThrThrAsp 980
Db 2913 TCAGAGCTCTCCGTACTGAAGAAATTACCGAAGCATGGGGCCGAGAGACCAACAGAC 2972
QY 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGlnLysAlaThrPro 1000
Db 2973 ATGGTGTCCGAGTTTCCACCTGACTGACTCCCAAGCACACAGAGAGAACCCCA 3032
QY 1001 ValGlnGlnValGlnSerGlyValLeuAspThrGlnGlnGlnLysArgGlnThrGlnAla 1020
Db 3033 GTTCAGAGGTAGAGAGTGTGTGTGATACAGAAAGAGAGAGCCGACAGCGAGCC 3092
QY 1021 IleLeuGlnAlaValAlaAspLysValLysGlnGlnSerGlnValProAlaThrGlnThr 1040
Db 3093 ATCTCCAAAGCCGTTGACAGACAAGGTGAAGAGAGTCCAGAGTCCGCAACCCAGACT 3152
QY 1041 ValGlnArgThrGlnSerLysAlaLeuGlnLysValGlnGlnValGlnLysAspSerGln 1060
Db 3153 GTGCAGAAACCGGGTCAAAAGACACTGAGAGAGTTGAGAGGTAGAGAGAGACTCCGAA 3212
QY 1061 ValLeuAlaSerGlnLysGlnLysAspValMetProLysGlyProValGlnGlnAlaGly 1080
Db 3213 GTCTGGCTTGGAGAAAGAGAGAGACTTATCCGAAAGAGACCCGTCCAGGAAGCTGGA 3272
QY 1081 AlaGlnHisLeuAlaGlnGlySerGlnThrGlnAlaThrProGlnSerLeuGlnVal 1100
Db 3273 GCTGAGCACTTTCACAGAGGCTGTGAGACTGAGACAGCTACTCCAGAGAGCTTGAATT 3332
QY 1101 ProGlnValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
Db 3333 CCTGAAGTCAACGCGAGATGTAGACCATGTCCGACAGTCCAGGTTATCAACTCCAGCAG 3392
QY 1121 LeuMetGlnGlnAlaValAlaProGlnSerSerGlnThrLeuThrAspSerGlnThrAsn 1140
Db 3393 CTGATGGAACAGGCGCTGGCCCTGATGATCATCCGAAACCTTGACAGACAGTGAACAAAT 3452
QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAlaAspGlnThrIle 1160
Db 3453 GGAACACTCCCTTATGACGATTACAGACACTGAGATGGGACACAGCAAGATGAAGAACTT 3512
QY 1161 AspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGlnGlnGln 1180
Db 3513 GACACCCAGGACAGTAAAGCCACTGCACTGTCAAGGCAAGTCAAGGTCAAGAAAGAG 3572
QY 1181 AlaAlaThrAlaGlnLysGlnGlnProSerThrLeuProAsnValProAlaGlnGln 1200
Db 3573 GCGGCTACTGCTCAGAAAGAGAGGAGCCCTTCACTACTTAATTAATGTCCAGCCAGGAA 3632
QY 1201 GlnHisGlyGlnGlnProGlyArgAspValLeuGlnProThrGlnGlnGlnLysThrAla 1220
Db 3633 GAAATGTGGGAAAGAACAGAGAGATGTTCTTGAACCTTACACAGCAAGAGATTACTGCT 3692
QY 1221 AlaAlaValProValLeuAlaLysThrGlnValGlyGlnGlnGlnValLysThrLeu 1240

Db 3593 GCACCGTCCCGCTTCTGGCAAGACTGAGGTGGTCAAGAGGGTGAGGTGAGTGTG 3752
 Qy 1241 AspGlyLeuLysValLysGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
 Db 3753 GATGGAGAAAAGTCAAGAAGAACAGAGGTGTTGTACACTCTGGACCCACAGTCAA 3812
 Qy 1261 LysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
 Db 3813 AAGCGTCTGATGTGACATATGACAGTGAAGTATGATGGAGTGGCGGTGTCAGGAAGA 3872
 Qy 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluMetGluThrAspValGlu 1300
 Db 3873 GAGAGTACTGAAGTGCAGAGTCTTAGCTGGAGGAGGAGATGGAACTGACGTGAA 3932
 Qy 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGluGlnGluThrAla 1320
 Db 3933 AAGGAGAAAAGGAGACACAAAGCCAGACCAAGTGAAGAGGTGAGCAGGAACAGCC 3992
 Qy 1321 AlaProGluHisGluGlyThrTyrGlyProValLeuThrLeuAspMetProSerSer 1340
 Db 3993 GCTCCTGAGCATGAAGAACCTACGGGAAGCCAGTCTGACACTTGACATGCCAGCTCA 4052
 Qy 1341 GluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLys 1360
 Db 4053 GAGAGGGGGAAGGACCTGGGAAGCCTTGGAGGAAGCCCTTCTCCAGACCAAGACAA 4112
 Qy 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
 Db 4113 GCAGGTGTCATAGAGTTCAGTTCAAGCTCAAGCCCTGGACACACAGTCACTCAACACGACGAA 4172
 Qy 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
 Db 4173 GCTGTGAAAAGGTCATAGAAAACGTTGTGATTTTCAGACAGCAGGTGAAGTCCAGAGTGT 4232
 Qy 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisThrThrLeu 1420
 Db 4233 GTAGGTGCACACTTATTACAGCTGAGAAGTCTCTCCAAAGCGGTGGCCACTGGACTCTT 4292
 Qy 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440
 Db 4293 CAGCATGCAGAGGACACGGTACCCCTGGGCTGAGTCTCAGGCAGAGATTCATCCCAATC 4352
 Qy 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAla 1460
 Db 4353 ATAGTAACCTCTGCTCTGAAAGCACCCTACCTCCTGACCTACAAAGGAGAAATAAGCGCA 4412
 Qy 1461 SerGlnArgGluArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
 Db 4413 TCCAGAGAGCGCATCAGAGGAGAGGACCAAGCCAGATGCTGCTGCTGATGCTGACGGC 4472
 Qy 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
 Db 4473 AAGGAGATGACAGCAATCGAAAAGTCTCTCAGGCTGAACTGAGATCTCGAACTTGAG 4532
 Qy 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
 Db 4533 AGTAAGACCAACAAGATTGCTGAACGTCATTACAGACAGCGGTGACCATTCGCACGT 4592
 Qy 1521 ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArg 1540
 Db 4593 ACAGAAAACGCCCCGAAACTCATGCTTATGATTTCAGACACCCAGGTTCTCTGCAATCGAG 4652
 Qy 1541 LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
 Db 4653 GCTGACAGCAGGAGCCCAACAGATGCTGGACAAAATGAAGTTGCCAAGATGAACAC 4712
 Qy 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGln 1580
 Db 4713 CCAGTGGCGCAGCCAGAGAGGACTTGCAGTCTGACCGTCTGAGGAGCATGGGCTCAG 4772
 Qy 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596

Db 4773 CCTCGAAATGCTTGGCGCGCTTGCAGTTGAAAGCCCGGTGTCAAAG 4820
 RESULT 7
 US-60-360-207-12210
 ; Sequence 12210, Application US/60360207
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig
 ; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
 ; FILE REFERENCE: CLO01321
 ; CURRENT APPLICATION NUMBER: US/60/360,207
 ; CURRENT FILING DATE: 2002-03-01
 ; NUMBER OF SEQ ID NOS: 47235
 ; SEQ ID NO 12210
 ; LENGTH: 6199
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; US-60-360-207-12210
 Alignment Scores:
 Pred. No.: 0 Length: 6199
 Score: 6663.00 Matches: 1357
 Percent Similarity: 88.90% Conservative: 76
 Best Local Similarity: 84.18% Mismatches: 156
 Query Match: 82.53% Indels: 23
 DB: 80 Gaps: 11
 US-09-902-432-4 (1-1596) x US-60-360-207-12210 (1-6199)
 Qy 1 MetGlyValAlaGlySerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
 Db 142 ATGGGTGTCAGGCGAGTTCACCGGACGAGCGGCCGCGAGCGCGGAGACGACG 201
 Qy 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
 Db 202 CCAGCGAGCTGGAGCTCAGTGGCCATGGCCGCGAGCGGAAAGCGTGGAGCGTGA 261
 Qy 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
 Db 262 GATCCCGCTGACCGGACCCCGCCACAGCTCCACAGAGAAATGTCAGCTGTCTGCC 321
 Qy 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluGlnGluGln 80
 Db 322 GTCAATGGTGTAGCTGAACAAGAGATGCCAGCTCCAGAGGAAAGCCAGGATGGCAA 381
 Qy 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
 Db 382 GAGGAGAGAGTCACTGTTGNAGATGTTGNACAGAGAGTCAAGAGTGTGAAGAAAA 441
 Qy 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
 Db 442 GACCGAGCTAAAGAAATGGCAGCCAGTTCACAGATTGTTGAAGATATCAAAAGGACGAG 501
 Qy 121 GlnGluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnValGluGlu 140
 Db 502 CAGGAGGAAACACCGGAAATAATCGAACACAGATCCCTGCTTCAGAGAGCAATGGGAGAA 561
 Qy 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
 Db 562 ATGGCGCAGGCTGCTGAGTCCCAGCTAATGAGTCCGCTTCAGAGAGGTATTAAATTT 621
 Qy 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
 Db 622 GTTGGTTTTAAATTACCGGTGAAGAGGATAAAAACGAAAGTCAGATACCGTCCAGCTA 681
 Qy 181 LeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
 Db 682 CTCACCTGTCAAGAGAGATGAAGCGGAGGCGGAGAGCCCTCCGTCGAGAGCAGGACAC 741
 Qy 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
 Db 742 CAAGAGCCCGGAGTGAGACCC---GTCCGCGAATCAGATCCAAAGAAAGTGAAGTGAAG 798
 Qy 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIlePro 240

QY 956 ThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAla 975
 DB 3007 ACAGCCGACAGAAACACAGAGGGCTCCGGCTCAAGAACTTACCGAAGCATCAGGGGCA 3066
 QY 976 GluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThr 995
 DB 3067 GAAGAGACCAACAGATGGTGTCTCGAGTTTCCAGCTGCCAGCTCCCGGACACACA 3126
 QY 996 GluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGlu 1015
 DB 3127 GAGGAAGCCACCCAGTTACGAGGAGTAGAGGGTGGCATGTAGATACGGAAGAACAGGAG 3186
 QY 1016 ArgGlnThrGlnAlaLeuGlnAlaValAlaAspLysValLysGluGluSerGlnVal 1035
 DB 3187 CGCCAGACGAGCGCGCTCCCAAGCGCTTGCACAAAGTGAAGAGGACTCCACAGGTG 3246
 QY 1036 ProAlaThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluVal 1055
 DB 3247 CTTGCAACCCAGACTCTGCAGAGAGCAGGCGCCGAAGCACTGGAGAAAGTGGAGGGTA 3306
 QY 1056 GluGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyPro 1075
 DB 3307 GAGGAGGACTCCGAGGTGCTGCTACCGAAGAGAGAGATGTTGTGCCGGAAGGACCC 3366
 QY 1076 ValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrPro 1095
 DB 3367 GTGCAGGAAGCTGAACTGAGCATCTTGCACAGGGCTCCGAGACTGTACAGGGCTACCCCA 3426
 QY 1096 GluSerLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnVal 1115
 DB 3427 GAGAGCCTTGAAGTTCTGAAGTACACAGAGGATGTAGCCGTCACCAATGCCAGGTT 3486
 QY 1116 IleLysLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThr 1135
 DB 3487 ATCAAGCACCAGAGCTGATGGAACAGGCTGTGCCCTCAGTCATCTGAAACCTTGACA 3546
 QY 1136 AspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln 1155
 DB 3547 GACAGTGAGACAAATGAAGTACTCCCTCCGAGATTTCAGACACTCCAAACGGGACACAG 3606
 QY 1156 GlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaValAlaArgGlnSerGln 1175
 DB 3607 CAAGACGAGCCGTGACAGCCAGGACAGTAATGCCATTGCCCGCTCAAGCAGTACAG 3666
 QY 1176 ValThrGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsn 1195
 DB 3667 GTCACTGAAGAGGAGGAGCTGTCTCAGACGGAGGGGCTTCAACA---CCATCTAGT 3723
 QY 1196 ValProAlaGlnGluHisGlyGluGluProGlyArgAspValLeuGluProThrGln 1215
 DB 3724 TTTCCAGCCAGGAAGAACACAGGAAACACAGGAAGGATGTTCTAGAACCCACA--- 3780
 QY 1216 GlnGluLeuThrAlaAlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGly 1235
 DB 3781 CAAGCGCTGGCTGCCGGGCGAGTCCCTATTCTGCAAGGCTGAGGTGGTCAAGGGT 3840
 QY 1236 GluValAspTrpLeuAspGlyGluLysValLysGlu-----GluGlnGlu 1250
 DB 3841 GAGCTGCCAGGTTTGTAGAGAAAGTCAAGACGGGACAGTGTGTAAAGAACTGGAG 3900
 QY 1251 ValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrThrAspSerGlu 1270
 DB 3901 GTGCTGTGCACACTGGACCCCAACAGTCAAAAGACTCTGACTTGACACGTCAGAGTGAA 3960
 QY 1271 ValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSerLeuSerLeu 1290
 DB 3961 GTATGAAGTGGCCAGATGTCAGGAACACTCAGAGTAATCAAGAACACAGACTATTAGCCCG 4020
 QY 1291 GluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGluGln 1310
 DB 4021 GAGAAAGAGAGATGGGAACCGAGCTTTGAAAGAGGAGAAACAGAGACCAAGACAGACAA 4080

QY 1311 ValSerGluGluGlyGluGlnGluThrAlaAlaProGluHisGluGlyThrThrGlyLys 1330
 DB 4081 GCCAGTGAAGAACATGAGCAGGAAACAGTGTCTCTGAGCATGAAGAAACCCCTTAAG 4140
 QY 1331 ProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGly 1350
 DB 4141 CAGTCTCTGACAGCTGACATGCTCTACTCAGAGAGGGGAAAGGCACTGGCAGCCTTGAA 4200
 QY 1351 GlySerProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSer 1370
 DB 4201 GGAAGCCCTTCTCTCCAGACCAAGCAAGCAGATTGCTAGAGGTTCAAGTTCAAAGC 4260
 QY 1371 LeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValLysGluThrValVal 1390
 DB 4261 TCAGACACACCACTGCTCAACCAACCGAAGTGTGCAAAAGGTCGAAAGAACTGTGGCA 4320
 QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
 DB 4321 ACTTCAGAGATGGATGAAAGTTTGGAGTGTGAGGTGCCAATCATATCCAGCTGAGAAG 4380
 QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
 DB 4381 CTCTCCGAAACCGTGGCTACGGGACTCTTCAGCATGGAGAGACACCGTGCCTCCAGGGG 4440
 QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
 DB 4441 CCTGAGTCTCAGCAGAGTCCATCCCAATATAGTAACCTCTCTCTGAAAGCATCCTA 4500
 QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluAsp 1470
 DB 4501 CATTCTGACCTTCAAGAGAAGTGGCGCATCCAGAAACAGAGATCAGATGAAGATAAC 4560
 QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
 DB 4561 AAGCCAGATGCTGCTCTGATGCTCGCGCAAGAGAGTGCACCAAGAGAGAAATCCTC 4620
 QY 1491 LysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIleValLeuAsnVal 1510
 DB 4621 AGGGCTGAACCTGAGATCTTGGAACTTGAGAGTAAGAGCAATAAGATTGTCCAGAGTGC 4580
 QY 1511 IleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHisAlaTyr 1530
 DB 4681 ATCCAGACAGCGCTGACCATGTTGACGTCACAGAAACAGCCCCGAAACCCACGCTTCT 4740
 QY 1531 AspSerGlnThrGlnValProAla-CysArgLeuAspSerArgGluProAsnArgCysTr 1550
 DB 4741 GATTACAGATCAGGTTCTCTGTGATGCGAGCTGACAGCCAGGAGGACACACAGATGCTG 4800
 QY 1550 pThrLysMetLys-----AspAlaLysMetLysHisProValProGlnPr 1565
 DB 4801 GACAAAGATGAAAGCGACCTTCAAGTCTCCCCCAAGATGGAACACTCAGTGCCTAGCC 4860
 QY 1565 OArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCysLe 1585
 DB 4861 CAGGAAGGACTTGGCGTTCTGATAGTCTGAGGATGAGCAAGGCTTCAGAAATGATC 4920
 QY 1585 uProArgLeuGlnLeuLysAlaProValSerLys 1596
 DB 4921 ACCACGCTTGCAGTTGAAAGTGCAGTGTCAAG 4954

RESULT 8
 US-08-978-277A-1
 ; Sequence 1. Application US/08978277A
 ; GENERAL INFORMATION:
 ; APPLICANT: Gelman, Irwin H.
 ; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Brumbaugh, Graves, Donohue & Raymond
 ; STREET: 30 Rockefeller Plaza
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA

```

ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,277A
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/665,401
FILING DATE: 18-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5074 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-978-277A-1

Alignment Scores:
Pred. No.: 6,1e+300 Length: 5074
Score: 6005.00 Matches: 1232
Percent Similarity: 97.33% Conservative: 7
Best Local Similarity: 96.79% Mismatches: 27
Query Match: 74.38% Indels: 14
DB: 13 Gaps: 4

US-09-902-432-4 (1-1596) x US-08-978-277A-1 (1-5074)
QY 329 GtUyStHgluPrOlaJaSeRgUngUngUProJaGlUaSPThRaSPGInaLaRg 348
DB 2 GAAAGACAGACCGACCTCGGAGGAGCAGCGCGGAGGAACACACAGCCAGCCAG 61
QY 349 LeuSeRaLaSPtYrGtUlySValGtUleuPrOleuGlUaSPGInaLaGlYasPleuGtU 368
DB 62 TTGTGAGCACTACACAGAGGTGAGGTGCTTGTGGAAGACCAAGGTGTGCTGAG 121
QY 369 AlaSeSeRglUglUuYcYsaLaProlaUaThRglUValPheAsPglUySMetGlu 388
DB 122 GATTCCTCAGAGAGAGAGTGTCTCTTGGCAACGAAAGTGTTAATGAGAAAGTGA 181
QY 389 AlaHSeGlUglUValaValaLaGlUValaHsValSeRThRaGlUyStHrgUngUngU 408
DB 182 GCCCACCAAGAGTTGTTCCAGAGGTCCAGCTGAGCACCGTGGAGAAACAGAGAGAG 241
QY 409 GtngUyGtngUyGtngUyGtngUyGtngUyGtngUyGtngUyGtngUyGtngUy 428
DB 242 CAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301
QY 429 ProPrOglUyLleuLaGlUuPProGlInglUValaProGlInglUaLaGlUuPProaLglUglU 448
DB 302 CCCCTGAGAAAGTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 361
QY 449 LeuMeLysSeRaRgUglUuMetCysValaSeRglYglYAsPHeStHrgInleuThRaSPleu 468
DB 362 CTGATGTAAGAGCAGAGAGAGTGTGTCTCTGAGGAGAGACCAACCTGAAGAGAGCTTA 421
QY 469 SerProGlUglUyStHrgUleuProLySHsProGlUglYlLeValaSeRglUValaGlUuMet 488

```

```

DB 422 AGTCTTGAAGAGAGACCGTCCCAACACCCAGAGGCAATTGTCACTGAGAGTGAAGT 481
QY 489 LeuSeSeRgInglUaRglLleYValGtngUySeRProleuYsLleuPheSeSeR 508
DB 482 CTGTCTCTCAGAAAGAAATCAAGTACAGGGAAGTCCCTTGAAGAACTCTTCAGTAGC 541
QY 509 SerGtYleuYsLleuSeRglYlYsLysGtngUySValRgUyGtngUyGtngUy 528
DB 542 TCAGGCTTAAAGAGCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
QY 529 GlUglUProGlYglUyTgInHsIleHsThRglUSeRProGlUSeRaLaSPGInaL 548
DB 602 GAAGAGCTTGAGAAATACCAACATTCACACCAATCCAGAGAGTGTGATAGAGAG 661
QY 549 LysGtngUySeSeRaLaSeSeRProGlUuPProGlUgUthRThCysLeuGlUyS 568
DB 662 AAGGAGAGAGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
QY 569 GlYProleuGlUaLaProGlInaSPglYglUaLaGlUglUglYThThSeRaSPglYglU 588
DB 722 GGGCCGCTGAGAGCACCCAG-GATGGGAGAGCTGAGAGAGAGAGAGAGAGAGAGAG 778
QY 589 LysLysARgUglUyLleThProTrPaLaSeRPhelYsLysMetValThProLySHs 608
DB 779 AAGAGAGAGAGAG--ATCACTCTCTGGGCAATCTTCAAAAGAGTGTACACCCAGAGAA 836
QY 609 ArgValaRgARgProSeRglUSeRAsPlySGtngUngUngUleuGtngUySValaLysSeRaLa 628
DB 837 CGGTC-CGAAAGCTTTCGAGAGTGCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 895
QY 629 ThRleuSeSeRThRaSPSeRThRaValaSeRglUuMetGlnaSPglUValaLysThRaGlY 648
DB 896 ACCTTGTCTCCACGATAGACACAGTGCAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 955
QY 649 GlUglUglUySProGlUglUProLySARgARgValaSPThRSeRaLaSPSeRProGlUaLa 668
DB 956 GAGGAAACAAAGCCAGAGAGAAACAAAGGCTGAGGAGTACTTCACTGTCTTGGAGAGCA 1015
QY 669 LeuLleYcYsValaGtngUySeSeRlYsLysARgLaARgUySAlaSeSeRSeRaSPglU 688
DB 1016 CTGATTTGTGTGATTCATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1075
QY 689 GlYglYProaRgThRleuGtngUyGtngUySeRHiSaRgLaGlUglUaLaSeRlYsAPly 708
DB 1076 ---GGGCAAGAGACACTGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1132
QY 708 sGlUaLaGlYThRaSPaLaValaProLaSeRThRgInglUglUaSPGInaLaGlYSe 728
DB 1133 AGAAGCCGA-ACAGAGCTGTGTCTGCAAGACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
QY 728 rSeSeRProGlUProaLaGlYSeRProSeRglUglUglUglUglUglUglUglUglU 748
DB 1192 TTCTTCACCCGAGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1251
QY 748 rPhelYsARgLeuValaThRProaRgUyLysSeRlYsLysSeRlYsLeuGlUyLysAla-- 767
DB 1252 ATTTAAAGATTAGTCACTCCAAAGAAAAAATCCAAAGTCAAAATCGGAAGAGAAAGA-AG 1310
QY 768 ----GluAsPSeSeRVal---GluGlnleuSeRThRgInglUglUglUglUglUglUgl 785
DB 1311 CCGAAGAGACTTACTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1370
QY 785 uSeRTrPaLaSeRlYleYsLysPheLleProGlYARgARgUyLysARgLaLaSPglYly 805
DB 1371 ACTTGGGTTTCATTAAAGAAATTCATCCCGGAGAGGAGAGAGAGAGAGAGAGAGAGAG 1430
QY 805 s-GtngUglUaLaThRaValaGlUaSPSeRglYProValaGlUleasngUlaSPaSPPro 825
DB 1431 GGCAAGAACAGACCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1490
QY 825 snValProLaLaValaProleuSeRglUyThRaSnLaValaGlUuRgUyLysMetGlu 845

```

Db 1491 ATGTCACGCGTCTGCTCTGCTGAGTATGATGTCAGTGGAGGAGGAGATCGAAG 1550
Qy 845 laGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuS 865
Db 1551 CCCAGGGGAATGCGGAGCTGCCAGCTGCTGGGCTGTGTA--GTGTCGAGAGCTCA 1607
Qy 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 895
Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCATGTGATGGACCCAGGGCAGTCACCA 1667
Qy 885 erValGluGluArgSerProSerTriPileSerAlaSerValThrGluProLeuGluHisT 905
Db 1668 GTGTCGAAGAGCGGTCTCTCTGTTGGATATCCGCTTCCTGTAACAGAACCTCTTGAACACA 1727
Qy 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAlaGluG 925
Db 1728 CAGCGGAGAGGACCATGCCACCTGTTGAAGAGGTCACTGAAAAAGACATCATTTGCGAAG 1787
Qy 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspMetValT 945
Db 1788 AAACCTCTGTGCTCACCCAGACGTTTACAGAGGTAAGATGCCCATGACGACATGGTCA 1847
Qy 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
Db 1848 CCAGTGAAGTGGATTTACCTCAGACGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCC 1907
Qy 965 rgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaV 985
Db 1908 GTACTGAAGAAGTTACCGAAGCATCGGGGCGGAGAGACCCACAGACATGGTGTCCGACG 1967
Qy 985 alSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005
Db 1968 TTTTCCAGCTGACTGACTCTCCACAGACCACAGAGGAAAGCCACCCAGTTTCAGAGGTAG 2027
Qy 1005 luSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaV 1025
Db 2028 AGGTTGTGTGCTAGATACAGAGAGAGAGGAGCCAGAGCCAGAGGCCATCTTCAAGCCG 2087
Qy 1025 alAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
Db 2088 TTGCAGACAAGGTGAAAGAGGAGTCCCGAGGTGCTGCTGCAACCCAGACTGTGCAGAGAACG 2147
Qy 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
Db 2148 GGTCAAAAGCACTGGAGAAGGTTGAGAGGTTAGAGGAGGACTCCGAAGTGTGCTGCTCGG 2207
Qy 1065 luLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085
Db 2208 AGAAAGAGAAGAGCTTATGCCGAAAGACCCGTCGAGGAAGCTGGAGCTGAGCATCTTG 2267
Qy 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
Db 2268 CACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGGCTTGAAGTTCTCTGAAGTCACAG 2327
Qy 1105 laAspValAspHisAlaAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnA 1125
Db 2328 CAGATGTAGACCATGTGCGCCAGCTGCCAGGTATTCAGAGCTCCAGCAGCTGATGGAACAGG 2387
Qy 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
Db 2388 CCGTGGCCCTTGAGTCTATCCGAACCTTGACAGACAGTGGACAAATGGGAAGCACTCCCT 2447
Qy 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAsps 1165
Db 2448 TAGCAGATTACAGACTGCAGATGGGACACAGCAGATGAACCATTTGACAGCCAGGACA 2507
Qy 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaG 1185
Db 2508 GTAAAGCCACTGCAGCTGTGAGGAGTTCAGAGGTTCAGAGGTCAGAGAAGAGGCGGCTACTGCTC 2567
Qy 1185 lnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluG 1205
Db 2568 AGAAAGAGGAGGCTTCCACACTACCTTAATAATGTTCCAGGCCAGGAAAGACATGGGGAAG 2627

Qy 1205 luProGlyArgAspValLeuLeuProThrGlnGlnGluLeuThrAlaAlaValProV 1225
Db 2628 AACCCAGGAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTGTCTGCGACCGCTCCCG 2687
Qy 1225 alLeuAlaLysThrGluValGlyGlnGluGlyValAspTrpLeuAspGlyGluLysV 1245
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGGTGAGGTGCTGCTGGTGGGAGAAAAAG 2747
Qy 1245 alLysGluGlnGlnValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
Db 2748 TCAAAAGAAACAGAGAGGTGTTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGCTGATG 2807
Qy 1265 alThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
Db 2808 TCACATATCACAGTGAAGTGTGGGAGTGGCGGTCTCAGGAAAAAGGAGAGTACTGAAG 2867
Qy 1285 alGlnSerLeuSerLeuGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
Db 2868 TGCAGAGTCTTTAGCCTGGAGGAGGAGATGGAATACTGACGTTGAAAGGAGGAAAAAGG 2927
Qy 1305 luThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAlaAlaProGluHisG 1325
Db 2928 AGACAAAGCCAGAGCAAGTGAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTCTGAGCATG 2987
Qy 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysA 1345
Db 2988 AAAGGAACCTACGGGAAGCCAGTCTGCACACTTGACATGCCAGTCCAGAGGAGGGAAGG 3047
Qy 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
Db 3048 CACTGGGAAGCCTTGAGGAAGCCCTTCTCTCCAGACCAAGCAAGCAGGTTGTCATAG 3107
Qy 1365 luValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysV 1385
Db 3108 AGGTTTCAGGTTCAAGCCTTGGACACAACTGACTCAACAGCAGGAGCTGTGGAAGAGG 3167
Qy 1385 alIleGluThrValValIleSerGluThrGlyLysSerProGluCysValGlyAlaHisL 1405
Db 3168 TCATGAACACGGTTGTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGCACACT 3227
Qy 1405 euLeuProAlaGluLysSerSerAlaThrGlyGlyHisThrThrLeuGlnHisAlaGluA 1425
Db 3228 TATTTACAGCTGAGAAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGAG 3287
Qy 1425 spThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProA 1445
Db 3288 ACACGGTATCCCTGGGGCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAATCTCTG 3347
Qy 1445 laProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluA 1465
Db 3348 CTCCTGAAAGCACCTTACATCTCTGACCTTACAGGAGAAATNAGCGCATCCACAGAGAGC 3407
Qy 1465 rgSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrA 1485
Db 3408 GATCAGAGGAAGAGGACAAAGCCAGATGCTGGTCTGATGCTGACGCAAGGAGGATACAG 3467
Qy 1485 laIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuSerLysSerAsnL 1505
Db 3468 CAATCAGAAAGTCTTCAAGGCTGAACCTTGAGATCTCTGGAATCTTGAGATTAAGGCAACA 3527
Qy 1505 ysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaP 1525
Db 3528 AGATTGTGCTGAACGTCATTCAGACAGCGCTTGACAGTTCGCACGTCACAGAAACAGCCC 3587
Qy 1525 roGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgG 1545
Db 3588 CCGAAACTCATGCTTATGATTCACAGACCCAGGTTCTCTGCAATGCGCTTGGACAGCAGGG 3647
Qy 1545 luProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHisProValProGlnP 1565
Db 3648 AGCCCAACAGATGCTGGAACAAAAATGAAAGTTGCCAAGATGAAACACCCAGTCCCGCAGC 3707

QY 1565 roatgcluaspleuglnvalleuthrvalleuglnualattpalaglnproarglyscyl 1585
 DB 3708 CCAGAGAGACTTGCAGAGCTCGACCGCTCTGAGGCAATGC-TCAGCTCGAAATGCT 3765
 QY 1585 euProArgLeuGlnLeuLysAlaProValSerLys 1596
 DB 3766 TCCCGCGCTTGCAGTTGAAAGCGCCGCTCTCAAG 3800
 RESULT 9
 PCT-US97-06830-1
 : Sequence 1, Application PC/TUS9706830
 : GENERAL INFORMATION:
 : APPLICANT: Gelman, Irwin H.
 : TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
 : STREET: 30 Rockefeller Plaza
 : CITY: New York
 : STATE: NY
 : COUNTRY: USA
 : ZIP: 10112-0228
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSeq Version 1.5
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US97/06830
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/665,401
 : FILING DATE: 18-JUN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Clark, Richard S
 : REGISTRATION NUMBER: 26,154
 : REFERENCE/DOCKET NUMBER: A30558 - 165/34008
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 212-408-2558
 : TELEFAX: 212-765-2519
 : TELEX:
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 5134 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA
 : HYPOTHEICAL: NO
 : ANTI-SENSE: NO
 : FRAGMENT TYPE:
 : ORIGINAL SOURCE:
 : PCT-US97-06830-1
 Alignment Scores:
 Pred. No.: 6,18e-300 Length: 5134
 Score: 6005.00 Matches: 1232
 Percent Similarity: 97.33% Conservative: 7
 Best Local Similarity: 96.78% Mismatches: 27
 Query Match: 74.38% Indels: 14
 DB: 1 Gaps: 4
 US-09-902-432-4 (1-1596) x PCT-US97-06830-1 (1-5134)
 QY 329 GtulyethrGluProAlaSerGluGlnGluProAlaGluAspThrAspGlnAlaArg 348
 DB 2 GAAAGACAGAGAGCACTCGAGAGCAAGAGCCGCGCAAGACACAGACCGAGCCAGG 61
 QY 349 LeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGluAspLeuGlu 368
 DB 62 TTGTGACGACACTTACGAGAGGTGAGAGCTGCTTTGGAGAGCAAGGTGTGTGACTGAG 121

QY 369 AlasSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
 DB 122 GCATGTAGAGAGAGAGATGTGCTCTTGGCAACGAAAGTGTGATGAGAGAGTGGAA 181
 QY 389 AlaHisGlnGluValAlaAlaGluValHisValSerThrValGluLysThrGluGlu 408
 DB 182 GCCACCAAGAAAGTGTGTCAGAGAGTCCAGTCAGACCCGTGAGAAAGACAGAGAGAG 241
 QY 409 GlnGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrGlyGluSerLeu 428
 DB 242 CAGGAGAGAGAGAGAGAGCTGAAGGGGCGTGTGTGAGAGAGAACAGGAAATCTCTTG 301
 QY 429 PropProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
 DB 302 CCCCCTGAGAAACTGGCTGAGCCCCAGAGAGTCCCCAGAAAGCTGAGCTGTGAGAG 361
 QY 449 LeuMetLysSerArgGluMetCysValSerGlyAlaPheHisThrGlnLeuThrAspLeu 468
 DB 362 CTGATGAAGAGAGAGAGATGTGTGTCTGTGAGAGAGACCACTCACTCACTGAGAGACTA 421
 QY 469 SerProGluGluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMet 488
 DB 422 AGTCTGAAGAGAGAGAGAGAGCTGCCCAACAGAGAGCATTTGTGAGAGTGAAGATG 481
 QY 489 LeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
 DB 482 CTGTCTCTCAGAAAGATCAAGGTACAGGAGAGTCCCTTGAAGAACTCTTCACTGATGC 541
 QY 509 SerGlyLeuLysLysLeuSerGlyLysLysGlyLysArgGlyGlyGlyGlyLysP 528
 DB 542 TCAGGCTTAAAGAGAGCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
 QY 529 GtugluProGlyGlyLysGlnHisIleHisThrGluSerProGluSerAlaAspGluGln 548
 DB 602 GAAGAGAGCTGAGAGATACCAACACATTCACACCGAATCCCAAGAGAGTGTGATGAGAG 661
 QY 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLys 568
 DB 662 AAGGAGAGAGAGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
 QY 569 GlyProLeuGluAlaProGlnAlaAspGlyGluAlaGluGluGlyThrThrSerAspGly 588
 DB 722 GGGCGGCTGGAAGACACCCAG-GATGGGAGAGCTGAGAGAGAGAGAGAGAGAGAGAG 778
 QY 589 LysLysArgGluGlyIleThrProThrAlaSerPheLysMetValThrProLysLys 608
 DB 779 AAGAGAGAGAGAG-ATCACTCCCTGGGAGATCTTCAAAAGATGTGTACACCCCAAGAA 836
 QY 609 ArgValArgArgProSerGluSerLysGluGluGluGluValLysSerAla 628
 DB 837 CGGTC-CGAAAGCTTCTGAGAGTACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 895
 QY 629 ThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGly 648
 DB 896 ACCTGTCTCTCCACTGATAGACACACTGTCAAGAAATGCAAGATGCAAAACTTTGGT 955
 QY 649 GtugluGlnLysProGluGluProLysArgArgValaAspThrSerValSerTProGluAla 668
 DB 956 GAGGAACCAAAAGCCAGAGAGAACCAAGAGGTGAGTACTTCACTGTCTTGGAGAGCA 1015
 QY 666 LeuIleCysValGlySerSerLysLysArgLysAlaArgLysAlaSerSerSerAspGlu 688
 DB 1016 CTGATTTGTGTGATCATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1075
 QY 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgLysGluGluAlaSerLysAspLys 708
 DB 1076 ---GGGCAAGAGACACTGGAGAGGGGAGACAGTACAGAGAGAGAGAGAGAGAGAG 1132
 QY 708 GtugluArgIleThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGly 728
 DB 1133 AGAGAGCGGA-ACAGAGCGTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
 QY 728 rSerSerProGluProAlaGlySerProSerGluGluGlyGlyValSerThrTProGluSe 748

Db 1192 TTCTCTACCCGAGCCAGCGGAAGCCCTTCCGAAGGGGAAGGTGTCTCCACTTGGGAGTC 1251
Qy rPheLysArgLeuValThrProArgLysLysSerLysLeuGluGluLysAla-- 767
Db 1252 ATTTAAAGATTAGTCACTCCAGAAAAAATCCAAAGTCAAACTGGAAGAGAAGA-AG 1310
Qy 768 ----GluAspSerVal---GluGlnLeuSerThrGluIleGluProSerArgGluGI 795
Db 1311 CCGGAAGGACTCTAGTTGTAGGAGCAGGTTGTCCACTGAGATCGAACCCGTTAGAGAAGA 1370
Qy 785 uSerTrpValSerIleLysLysPheIleProGlyValArgLysLysLeuArgAlaAspGlyLY 805
Db 1371 ATCTTGGGTTTCCATTAGAAATTCATCCCGGAGCGGCGGAAGAAAGGCGAGATGGGA 1430
Qy 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspProA 825
Db 1431 GGCAAGAAACAAGCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGACGAGCCTG 1490
Qy 825 snValProAlaValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluA 845
Db 1491 ATGTCCAGCGCGTGGTCTCTGTCTAGTATATGTCAGTGGAGGAGGAAGATGGAAG 1550
Qy 845 laGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLys 865
Db 1551 CCAGGGGAATGGGAGCTGCCAGCTGCTGGGGCTGTGTA---GTGTCCGAGAGCTCA 1607
Qy 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTCCAGTTCATTGATGGGACCAGGCGAGTCA 1667
Qy 885 erValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisT 905
Db 1668 GTGTGGAAGACGGGTCTCTTCGTGGATATCCGCTTCCGTAACAGAACCTCTTTGAACACA 1727
Qy 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAlaGluG 925
Db 1728 CACGGGAAGGACATGCCACCTGTTGAGAGGTCACTGAAAGAGACATCATTTGCGAAG 1787
Qy 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspMetValT 945
Db 1788 AAACCTCTGTGCTCACCCAGACGTATCCAGAGGTAAAGATGCCATGACACATGGTCA 1847
Qy 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
Db 1848 CCAGTGAAGTGGATTTTCACTCTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCC 1907
Qy 965 rgThrGluGluValThrGluAlaSerClyAlaGluGluThrThrAspMetValSerAlav 985
Db 1908 GTACTGAAGAAGTTTACCGAAGCATCGGGGCGGAAGAGACCAGACATGGTGTCCGAG 1967
Qy 985 alSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005
Db 1968 TTTCCAGCTGACTGACTCTCCACAGACACACAGAGGAGGCCACCCAGTTTCAGAGGTAG 2027
Qy 1005 luSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAlav 1025
Db 2028 AGGGTGGTGTGTAGATACAGAAGAAGAGAGAGCGCCACAGCGCAGGCCCATCTCCAAGCCG 2087
Qy 1025 alaAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
Db 2088 TTGCAGACAAGGTGAAGAGAGGTCCAGGTGCGCTGCAACCCAGACTGTGCAGAGAACGG 2147
Qy 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
Db 2148 GGTCAAAAGCACTGGAGAAGTTGAGAGGTAGAGGAGGACTCCGAAGTGTGCTTCGG 2207
Qy 1065 luLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085
Db 2208 AGAAAGAGAAGGACGTTATGCCGAAAGGACCCGTCAGGAAGCTGGAGCTGAGCATCTTG 2267
Qy 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105

Db 2268 CACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCACAG 2327
Qy 1105 laAspValAspHisValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGlnA 1125
Db 2328 CAGATGTAGACCATGTGCGCACAGTGCAGGTTATCAAGCTCCAGCAGCTGTATGTAACAGG 2387
Qy 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
Db 2388 CCGTGGCCCTCAGTGCATCCGAACCTTGACAGACAGTGAAGCAATGGAAGCAGCTCCCT 2447
Qy 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspS 1165
Db 2448 TAGCAGATTACAGCACTGCAGATGGACACAGCAAGATGAACCATTTGACAGCAGGACA 2507
Qy 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaG 1185
Db 2508 GTAAAGCCACTGCAGCTGTGAGCAGCTACAGGTACAGAAAGAGGCGGCTACTGCTC 2567
Qy 1185 luLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluG 1205
Db 2568 AGAAAGAGAGGCGCTTCGACACTACCTTAATGTTCAGGCCAGGAAGCAATGGGGAAG 2627
Qy 1205 luProGlyArgAspValLeuGluProThrGlnGlnLeuLeuThrAlaAlaValProV 1225
Db 2628 AACCAGAAAGAGATGTTCTTGAACCTACACAGCAGAGCTTGTCTGTCGACCGCTGCCG 2687
Qy 1225 alLeuAlaLysThrGluValGlyGlnGluGlyValAspTrpLeuAspGlyGluLysV 1245
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGTGAGGTTCGACTGGTTGGATGGGAAAG 2747
Qy 1245 alLysGluGlnGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
Db 2748 TCAAGAAAGAACAGGAGGTGTTGTACACTTGGACCCACAGTCAAAAGGCTGCTGATG 2807
Qy 1265 alThrTrpAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
Db 2808 TGACATATACACAGTGAAGTGAAGTGGAGTGGCGGTGTTCAGAAAGAGGAGTACTGAAG 2867
Qy 1285 alGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
Db 2868 TGCAGAGTCTTAGCCTGGAGGGAGAGATGGAACCTGACGTTGAAAAGGAGGAAAGGG 2927
Qy 1305 luThrLysProGluGlnValSerGluGluGlyGlnGlnThrAlaAlaProGluHiG 1325
Db 2928 AGACAAAGCCAGACGAAGTGAAGTGAAGAGGTGAGCAGGAAACAGCGCTCTCTGAGCATG 2987
Qy 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluAtqGlyLysA 1345
Db 2988 AAGGAACCTACGGGAAGCCAGTCTTGACACTTGACATGCCAGCTCAGAGAGGGGAAGG 3047
Qy 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
Db 3048 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGACAGAGGTTGCATAG 3107
Qy 1365 luValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysV 1385
Db 3108 AGGTTCAAGTTCAAAGCCTGGACACAACTGACTCAAAACAGCAGAAAGCTGTGGAAAGG 3167
Qy 1385 alIleGluThrValValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisL 1405
Db 3168 TCATGAAGACGGTTGTGATTTTCAGACAGAGGTGAAGTTCAGAGTGTGTAGGTGCACACT 3227
Qy 1405 euLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluA 1425
Db 3228 TATTACAGCTGAGAAGTCTCTGCAACGGGTGGCCACTTGGACTCTTCAGCATGCGAGG 3287
Qy 1425 spThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleValThrProA 1445
Db 3288 ACACGGTACCCCTGGGGCCTGAGTCTCAGCGAGAATCCATCCCAATCATAGTAACCTCTG 3347
Qy 1445 laProGluSerThrLeuHisProAspLeuGlnGlyLysLeuSerAlaSerGlnArgGluA 1465
Db 3348 CTCCTGAAAGCACCCCTACATCTCTGACCTCAAGAGAGAAATAAGCGCATCCCGAGAGAGC 3407

QY 629 ThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValIysThrValGly 648
Db 896 ACCTGTGCTCCACTGATAGCAGGTGTGAGAAATGCAAGATGAAGTCAAAATCTGTGGT 955
QY 649 GluGluGlnLysProGluGluProLysArgArgValAspThrSerValSerTrpGluAla 668
Db 956 GAGGACAAAAGCCAGGAGAACCAAGCGTAGGGTGATCTTCACTGTCTTGGGAAGCA 1015
QY 669 LeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGlu 688
Db 1016 CTGATTGTGTGCGATCATCCAGAGAGAGCAAGGAAGCATCTCTTCAGATATAAGA 1075
QY 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgAlaGluGlnAlaSerLysAspLys 708
Db 1076 ---GGGCAAGGACACTGGAGGGGGACGTCAAGAGCAGAGAGGAGCCCAAGACAA 1132
QY 708 sGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySe 728
Db 1133 AGAAGCCGA-ACAGACGCTGTCTCTGCAGACACCCAGGACAGGACCCAGCGAAGAG 1191
QY 728 rSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSe 748
Db 1192 TTCCTCACCCGAGCCAGCGGAAGCCCTCCGAAGGGGAAGGTGTCTCCACTTGGGAGTC 1251
QY 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGlnLysAla-- 767
Db 1252 ATTTAAAGATTAGTCTACTCCAGAAATAATCCAACTCAAACTGGAAAGAGAAGA-AG 1310
QY 768 ---GluAspSerSerVal---GluGlnLeuSerThrGluIleGluProSerArgGluG 785
Db 1311 CCGAAGGACTCTAGTTGTAGGAGCAGGTTGTCCACTGAGATCGAACCGTGTAGAGAAGA 1370
QY 785 uSerTrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLys 805
Db 1371 ATCTTGGTTTCCATTAAAGAAATTCATCCCGGAGCGCGAAGAAAGGCGCAGATGGAA 1430
QY 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProA 825
Db 1431 GGAAGAACAAGCCACTGTGGAGAGACTCAGGGCCAGTGGAGATAAATAGAGACGCGCTG 1490
QY 825 snValProAlaValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluA 845
Db 1491 ATGTCCAGCGCTGCTGCTCTGTCTGAGTATGATGCAGTGGAGAGGAGAAAGATGGAG 1550
QY 845 laGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuS 865
Db 1551 CCAGGGGAATGCGGAGCTGCCAGCTGCTGGGCTGTGTA---GTGTCCGAGGAGCTCA 1607
QY 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGGACCCAGGCGAGTCA 1667
QY 885 erValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisT 905
Db 1668 GTGTGGAAGAGCGGTCTCTTCTGGTGATATCCGCTTCCGTAACAGAACCTCTTGAACACA 1727
QY 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAlaGluG 925
Db 1728 CAGCGGAGAAAGCCATGCCACCTGTTGAGAGGTCACTGAAAGAGACATCATTTGCAGAAG 1787
QY 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945
Db 1788 AAACCTCTGTGCTCACCCAGACGTTACAGAGGGTAAGATGCCCATGACGACATGTGCA 1847
QY 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
Db 1848 CCAGTGAAGTGGATTTTCACCTCAGAAAGCTGTGACAGCCACAGAGACCTTCAGAGCTCTCC 1907
QY 965 rgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlav 985
Db 1908 GTACTGAAGAAAGTTACCGAAGCATCGGGGCGGAGAGACACAGACATGGTGTCCCGAG 1967
QY 985 alSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005

Db 1968 TTTCCAGCTGAGTGTACTCCAGAGACACACAGAGAGAAAGCCCCAGTTCAGGAGGTAG 2027
QY 1005 luSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlav 1025
Db 2028 AGGGTGGTGTCTAGATACAGAAGAAGAGAGCGCCAGAGCGCCATCTCTCAAGCCG 2087
QY 1025 alAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
Db 2088 TTGCACACAAGGTGAAAGAGAGGTCCAGGTGCTCTCAACCCAGAGCTGTGCAGAGAACGG 2147
QY 1045 lySerLysAlaLeuGluLysValGluGluValIleLysLeuAspSerGluValLeuAlaSerG 1065
Db 2148 GGTCAAAAGCCTGGAGAAAGGTTGAGAGGTAGAGAGGACTCCGAAGTGTCTGGCTTCGG 2207
QY 1065 lLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085
Db 2208 AGAAAGAGAAAGACGTTATGCCGAAAGACCGGTGCAGGAGCTGGAGCTGAGCATCTTG 2267
QY 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
Db 2268 CACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTTCCTGAAGTCACAG 2327
QY 1105 laAspValAspHisValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGlnA 1125
Db 2328 CAGATGTAGACCATGTGCCACGTGCCAGGTATCAAGCTCCAGCAGCTGATGGAAACAG 2387
QY 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
Db 2388 CCGTGGCCCTTGAGTCTATCCGAAACCTTCACAGACAGTGCAGACAAATGGAGCATCTCC 2447
QY 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspS 1165
Db 2448 TAGCAGATTTCAGACACTGCAGATGGGACACAGCAAGATGAACCAATTGACAGCAGGACA 2507
QY 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluAlaAlaThrAlaG 1185
Db 2508 GTAAAGCCCACTGCAGCTGTGCAGCAGTCAAGAGGAGGAGGCGGTACTGCTC 2567
QY 1185 lnLysGluGluProSerThrLeuProAsnValProAlaGlnGluGluHisGlyGluG 1205
Db 2568 AGAAAGAGAGGCTTCGACACTACCTAATATGTTCCAGCCAGGAGAAACATGGGGAAG 2627
QY 1205 luProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProV 1225
Db 2628 AACCGGAAGAGATGTTCTTGAACCTTACACAGCAGAGCTTGTCTGTCAGCGCTGCCCG 2687
QY 1225 alLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysV 1245
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGGGTGAGGTTGACTGGTTGGATGGAGAAAAG 2747
QY 1245 alLysGluGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
Db 2748 TCAGAAAGAACAGAGAGGTGTTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGTGATG 2807
QY 1265 alThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
Db 2808 TGACATATGACAGTGAAGTGAAGTGGGAGTGGCGGGTGTCCAGGAAAAGAGAGTACTGAAG 2867
QY 1285 alGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
Db 2868 TGCAGAGTCTTAGCTCTGAGAGGAGAGATGGAATCTGACGTTGAAAAGGAGAGAAAAGG 2927
QY 1305 luThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAlaAlaProGluHisG 1325
Db 2928 AGCAAAAGCCAGAGCAAGTGAAGAGAGGTGAGCAGGAAACAGCCGCTCTCTGAGCATG 2987
QY 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluAlaTrpGlyLysA 1345
Db 2988 AANGGAACCTACGGGAAGCCAGTCTCTGACACTTGACATGCCAGCTCAGAGAGGGGGAAG 3047
QY 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365


```

Db 3048 CACTGGGAAGCCTTGGAGGAGCCCTTCTCCACAGCAACCAAGAGGTTGCTAG 3107
Qy 1365 JuvAIGlnValIGlnSerLeuAspThrValThrGlnThrAlaGluValGluVal 1385
Db 3108 AGGTTCAAGTTCAAAGCCCTGGACACAAACAGTCACTCAACACAGAAAGCTGGGAAAG 3167
Qy 1385 AllIGlnThrValValIleSerGluThrGluGluSerProGluCysValAlaHisLe 1405
Db 3168 TCATAGAAACGGGTGTGATTTCAGAGACAGGTGAAGTCCAGATGTGTAGGTGCACACT 3227
Qy 1405 euleProAlaGluLysSerSerAlaThrGlyYhiStrpThrLeuGlnHisAlaGlu 1425
Db 3228 TATTACAGCTAGAGAGTCTCTGGCAAGGGGTGGCCACTTTCAGCATGAGAGAG 3287
Qy 1425 sPthValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrPro 1445
Db 3288 ACACCGTACCCCTGGGGCTGAGTCTCAGGCAAGATCCCAATCCCAATCATGTAACCTCTG 3347
Qy 1445 lAProGluSerThrLeuHisProAspLeuGlnGluIleSerAlaSerGlnArgGlu 1465
Db 3348 CTCTGGAAGGCAACCTTACATCTGACTACAGAGAAATTAAGCGATCCCAAGAGAGC 3407
Qy 1465 rSergLugluLysAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThra 1485
Db 3408 GATTCAGAGAGAGAGAGCAAGCCAGATGCTGTCTGATGCTGACGGCAAGAGAGTACAG 3467
Qy 1485 lAlIeGluLysValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsn 1505
Db 3468 CAATCGACAAAGTCTCCAGAGCTGAACCTGAGATCTGGAACCTTGAGATTAAGCAACA 3527
Qy 1505 yAlIeValLeuAsnValIleGlnThrAlaValAspGlnIleAspGlnThrGluThrAlaP 1525
Db 3528 AAGTGTGTCTGAACGATTCACAGACCGGTGACCAAGTTCGACCTACAGAAACAGCCC 3587
Qy 1525 rGluThrHisAlaTrpAspSerGlnThrGlnValProAlaCysArgLeuAspSerArg 1545
Db 3588 CCGAAATCATGCTTATGATTTCACAGACCCAGGTTCTGCAATGGCTTGAGACAGAGG 3647
Qy 1545 lUpProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHisProValProGlnP 1565
Db 3648 ACCCCCAACAGATGCTGGAACAAATGAATGAAGTTGCCAAGATGAACACCCACAGTCCCAAC 3707
Qy 1565 rAArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCysL 1585
Db 3708 CCAGAGAGAGACTTGCAAGTCTGACCGTCTGAGGCAATGGC--TCAGCTCGGAAATGCT 3765
Qy 1585 euProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 3766 TGCCCGGCTTGCAAGTGAAGCCCGGTGTCAAG 3800

RESULT 11
US-09-902-432-1
; Sequence 1, Application US/09902432
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A3058-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5134
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-1

```

```

Alignment Scores:
Pred. No.: 6,18e-300 Length: 5134
Score: 6005.00 Matches: 1232
Percent Similarity: 97.33% Conservative: 7
Best Local Similarity: 96.78% Mismatches: 27
Query Match: 74.38% Indels: 14
DB: 34 Gaps: 4

US-09-902-432-4 (1-1596) x US-09-902-432-1 (1-5134)
Qy 329 GluYsThrGluProAlaSerGluGluGlnLupProAlaGluAspThrAspGlnAlaArg 348
Db 2 GAAAGACAGAGCCAGCTCCGAGGAGCAGAGCCGCGAGAAAGACACAGACAGCCACG 61
Qy 349 LeuSerAlaAspTrpGlnLysValGluLeuProLeuGluAspGlnValGlyAspLeu 368
Db 62 TTGTACGACAGACTACGAAAGAGTGGAGCTGCTTGGAAAGACCAAGTTGTGACCTGGAG 121
Qy 369 AlAserSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
Db 122 GCATGTCAGAGAGAGAGAGTGTGCTCTTGGCAACGAGAGTGTGTGATGAGAAAGTGA 181
Qy 389 AlAsISGlnGluValAlaValAlaGluValHisValSerThrValGluLysThrGluGlu 408
Db 182 GCCCACCAAGAAAGTTGTTGACAGAGTCCACGTGACACCTGGAGAGACACAGAGAGAG 241
Qy 409 GlnGlyGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrGlyGluSerLeu 428
Db 242 CAGGAGAGAGAGAGAGAGCTGGAAGGGGGCTGGTGTGAGAAAGAACGAGAAATCCTTG 301
Qy 429 ProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
Db 302 CCCCTGAGAACTGGCTGAGCCCCAGAGAGTCCCCAGAAAGCTGAGACCTGTGAGAG 361
Qy 449 LeuMetLysSerArgGluMetCysValSerGlyGlyAspHisStrpGlnLeuThrAspLeu 468
Db 362 CTGATGAAGAGAGAGAGATGTGTCTCTGAGAGAGACCACTCAACCTGACAGACCTA 421
Qy 469 SerProGluGluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMet 488
Db 422 AGTCTTGAAGAGAGAGAGACCTGCCCCCAACACCAAGAGCATTTGACAGAGTGAAGATG 481
Qy 489 LeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
Db 482 CTGTCTCTCAGAGAAAGATCAAGGTACAGGAACTCCCTTGAAGAACTCTTCAGTATGC 541
Qy 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyValYasp 528
Db 542 TCAGGCTTAAAGAGACTGTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
Qy 529 GlnGluProGlyGluTrpGlnHisIleHisThrGluSerProGluSerAlaAspGluGln 548
Db 602 GAAAGAGCTCGAGAGATTAACAACATTCACACCGAATCCCAAGAGAGTGTGATGAGAG 661
Qy 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrCysLeuGluLys 568
Db 662 AAGGAGAGAGAGTCTGCGTCTGCCCCCAGAGAGCTGAGAGAGACCACTGTCTGAGAAA 721
Qy 569 GlyProLeuGluAlaProGluAspGlyGluAlaGluGluGluGlyThrThrSerAspGly 588
Db 722 GGGCCGCTGGAAGACCCAG--GATGGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 778
Qy 589 LysLysArgGluGlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLys 608
Db 779 AAGAGAGAGAGAG--ATCATCTCCCTGGGCATCTTCAAAAAGATGTGACACCCAGAAA 836
Qy 609 ArgValArgArgProSerGluSerAspLysGluGluGluGluValLysSerAla 628
Db 837 CGGTC--CGAAGACTTTCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 895
Qy 629 ThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGly 648

```

Db 896 ACCTTGCTCCACTGATAGCAGTGTGAGAAATGCAAGATGAAGTCAAACTGTTGGT 955
Qy 649 GluGluGlnLysProGluGluProLysArgValAspThrSerValSerTrpGluAla 668
Db 956 GAGGAACAAAAGCCAGAGGAACCAAGCGTAGGGTGATATTCAGTGTCTTGGGAAGCA 1015
Qy 669 LeuLeuCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGlu 698
Db 1016 CTGATTGTGTGCGGATCATCCAGAGAGAGCAAGGAAGGCATCTCTTCAGATATAAGA 1075
Qy 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgAlaGluGluAlaSerLysAspLys 708
Db 1076 ---GGGCCCAAGACACTGGGAGGGGACAGTCAAGAGCAGAGAGGCCCAAGAACAA 1132
Qy 708 sGluAlaGlyThrAspAlaValProLaserThrGlnGluInAspGlnAlaGlnGlySe 728
Db 1133 AGAAGCCGA-ACAGACGCTGTTCTGCCAGCACCCAGGAGCAGGACCAAGCGCAAGGAAG 1191
Qy 728 rSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSe 748
Db 1192 TTCTCTCACCCGAGCCAGCGGAAGCCCTTCCGAAGGGGAGGTGTCTCCACTTTGGGAGTC 1251
Qy 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGluLysAla-- 767
Db 1252 ATTTAAAGATTAGTCACTCCAAAGAAAATAATCCAAGTCAAACTCGAAGAGAAAGA-AG 1310
Qy 768 ----GluAspSerSerVal----GluGlnLeuSerThrGluLeuGluProSerArgGlu 785
Db 1311 CCGGAAGGACTCTAGTTGTAGGAGCAGGTGTCCACTGAGATCGAACCCGTGTAGAGAGA 1370
Qy 785 uSerTrpValSerIleLysLysPheLysProGlyArgArgLysLysArgAlaAspGlyLys 805
Db 1371 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGGAGCGCGGAAGAAAGGCGACATGGGAA 1430
Qy 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProA 825
Db 1431 GGCAAGAACAAAGCCACTGTGGAAAGACTCAGGGCCAGTGGAGATAAATGAGGAGCGCTG 1490
Qy 825 snValProAlaValAlProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluA 845
Db 1491 ATGTCCACCGCGTGTCTCTCTGTCTGTGATGATGATGATGATGATGATGATGATGATG 1550
Qy 845 laGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuS 865
Db 1551 CCCAGGGGAATGCGGAGCTGCCAGCTGTCTGGGGCTGTGTA---GTGTCCGAGGAGCTCA 1607
Qy 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGGACAGGGCAGTCACCA 1667
Qy 885 erValGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHist 905
Db 1668 GTGTCAAGAGCGGTCTCTCTCGTGGATATCCGTTCCGTAACAGAACCTCTTGAACACA 1727
Qy 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAlaGluG 925
Db 1728 CAGCGGAGAAAGCATGCCACCTGTTGAAGAGGTCACTGAAAAGACATCATTTGCAGAAAG 1787
Qy 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945
Db 1788 AAACTCTGTGCTCACCCAGACGTACCAGAGGGTAAAGATGCCCATGACGACATGGTCA 1847
Qy 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
Db 1848 CCAAGTGAAGTGGATTTCCTCTGAGAGTGTGTACAGCCACAGAGACCTCAGAGGCTCTCC 1907
Qy 965 rgThrGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaVal 985
Db 1908 GTACTGAAGAAGTTACCGAAGCATCGGGGCCGGAAGAGACCAAGAGACATGCTGTCCGCGAG 1967
Qy 985 aIsrGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005
Db 1968 TTTTCCAGCTGACTGACTCCCCAGACACCAAGAGGAAGCCACCCAGGTTTCAGAGGTAG 2027

Qy 1005 luSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaVal 1025
Db 2028 AGGGTGTGTCTAGATACAGAAAGAGAGCGCCAGAGCGGCACTCTCTCCAGCGCG 2087
Qy 1025 alAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrGlnAlaArgThrG 1045
Db 2088 TTGCAGACAAAGGTGAAAGAGAGTCCCAGGTGCTGCAACCCAGAGCTGTGCGAGAAACGG 2147
Qy 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
Db 2148 GGTCAAAAGCACTGGAGAAGGTTGAGGAGGTAGAGGAGARCTCCGAGTGTCTGGCTTCGG 2207
Qy 1065 luLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085
Db 2208 AGAAAGAGAAGACGCTTATGCCGAAAGGACCCGTGCGAGGAAGCTGGAGCTGAGCATCTTG 2267
Qy 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
Db 2268 CACAGGGCTCTCAGACTGGACAGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCACAG 2327
Qy 1105 laAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnA 1125
Db 2328 CAGATGTAGACCATGTGCGCACGTGCCAGGTTATCAAGCTCCAGCAGCTGATGGAAACAGG 2387
Qy 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
Db 2388 CCGTGGCCCTCGAGTCATCCGAAACCTTGACAGACAGTGCAGCAAAATGGAAGCACTCCCT 2447
Qy 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAsps 1165
Db 2448 TAGCAGATTTCAGACACTGCAGATGGGACACAGCAAGATGAAACCATTTGACAGCAGACA 2507
Qy 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluAlaAlaThrAlaG 1185
Db 2508 GTAAGCCACTGCAGCTGTGAGGAGTACAGGTACAGAAAGAGGGCGGTACTGTCTC 2567
Qy 1185 lnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluG 1205
Db 2568 AGAAAGAGGAGCCTTCGACACTACCTAATATGTTCCAGCCCGAGGAAGCAATGGGGAAG 2627
Qy 1205 luProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProV 1225
Db 2628 AACCCAGGAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTGTCTGTGCGCGCTGCCCG 2687
Qy 1225 alLeuAlaLysThrGluValGluGlnGluGluValAspTrpLeuAspGlyGluLysV 1245
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGTGAGGTTGACTGTTGGATGGAGAAAAAG 2747
Qy 1245 alLysGluGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
Db 2748 TCAAAAGAAACAGGAGGTGTTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGCTGATG 2807
Qy 1265 alThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
Db 2808 TGACATATACAGTCAAGTGAAGTGGAGTGGCGGGTGTCCAGGAAAAGGAGAGTACTGAAG 2867
Qy 1285 alGlnSerLeuSerLeuGluGluGluMetGluThrAspValGluLysGluLysArgG 1305
Db 2868 TGACAGTCTTACCTCGAGGAGGAGAGATGGAACCTGACCGTTGAAAAGGAGAAAAAGG 2927
Qy 1305 luThrLysProGluGlnValSerGluGluGluGlnGluThrAlaAlaProGluHisG 1325
Db 2928 AGCAAAAGCCAGACAGTGAAGAGGTGAGCAGGAACACAGCCGCTCTCTGAGCATG 2987
Qy 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysA 1345
Db 2988 AAAGGAATACCGGAAGCCAGTCTCTGACACTTGATGCCAGTCCAGAGAGGGGGAAGG 3047
Qy 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
Db 3048 CACTGGGAAGCCTTGAGGAGGCCCTTCTCTCCAGACCAAGCAAGCAGGTTGCATAG 3107

QY 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
Db 1185 GTAGACAGAGAAGAACGGAAGGCGAGGTTCCCTCCGAGAAACTGACGCGCTCCGAG 1244
QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
Db 1245 CAAGCCACCACAGGAGCGGAGAGTCCACAGCGCCCGTTATCAGCTGAATAT 1304
QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1305 GAGAAAGTTGAGCTGCCCTCAGAGGAGCAAGTCACTAGTGCCTCGCAGGACCTTCTGAAGAG 1364
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln--Glu 392
Db 1365 AAACCTGCTCCGTGGCGAGAGAGTGTTCATGAGAAATGAGAGTCCACCAAGAGAG 1424
QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGly 412
Db 1425 GTTGTGGCCGAGTCCACGTCAGCACCGTGGAGGAGAGAACCGNAGAGCAG----- 1475
QY 413 GlyGluAlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1476 -----AAAAAGAGGTGGAAGAACACAGCAGGCTGTGTGCCAGCTGAAGAA 1520
QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSer 452
Db 1521 TTGTTGGAATGGATGATGAGAACCTTCAGGAAGCCGAACTGCCAAGGAGCTGGTGAAGCTC 1580
QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1581 AAAGAAACGTGTGTTCCGAGAGGACCTTACAGGGAGCTGACCTCAGTCTGTATGAG 1640
QY 473 LysThrLeuProLysHisProGluGlyLeValSerGluValGluMetLeuSerSerGln 492
Db 1641 AAGTGTGTCCAAACCCCGAAGGCGTGTGTGAGTGGGAAATGCTGTCAACAG 1700
QY 493 GluArgLeuLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
Db 1701 GAGAGAAATGAAGTGCAGGGAAGTCCATAAGAAAGCTTTTACAGCACTGGCTTAAAA 1760
QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGly 532
Db 1761 AGCTTTCTGGAAAGAAACAGAAAGGGAAGA--GGAGAGGAGACGAGGAATCAGGG 1817
QY 533 GluTyrGlnHisLeHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1818 GAGCACACTCAGGTTCAGCGGATCTCCGACAGCAGGAGGAGCAAAAGGGCGAGAGC 1877
QY 553 SerAlaSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1878 TCTGCCCTCATCCCTGAGGAGCCGAGGAGATCACGTGTCTGAAAAAGGGCTTAGCCGAG 1937
QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1938 GTGCAGCAGGATGGGAAGCTGAAGAGGAGTACTTCCGATGGAGAGAAAGAAAGAGAA 1997
QY 593 GlyLeuThrProTyrAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
Db 1998 GGTGTCACTCCCTGGGCATCATTCAAAAGATGTGTGAGCCCAAGAGCGGTGTAGACGG 2057
QY 613 ProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
Db 2058 CCTTCGAAAGTATATAAGAGAGATGAGCTGACAAAGGTCAAGAGCGCTACCTGTCTTCC 2117
QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluLys 652
Db 2118 ACCGAGACACAGCCTCTGAATTCAGAGAAATGAAGGGAGCGGTGGAGAGCCAAAG 2177
QY 653 ProGluGluProLysArgValAspThrSerValSerTyrGluAlaLeuLeuLysCysVal 672
Db 2178 CCGAAGAACCAAGCCAGGTGATACCTCAGTATCTTGGGAGGCTTAATTTGTGTG 2237
QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg 692

Db 2238 GGATCATCCAAAGAAAAGACAGAGAGGGTCTCTCTGTATGAGGAAGGGGACCAAAA 2297
QY 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2298 GCAATGGAGGAGACCCACCAGAAAGCTGATGAGCCCGGAAACACAAAGAGACGGGACA 2357
QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerProGlu 732
Db 2358 GACGGGATCTTCTGTTTCCCAAGAACATGATCCAGGGCAGGGAAGTCTCTCCCGGAG 2417
QY 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTyrGluSerPheLysArgLeu 752
Db 2418 CAAGCTGGAGAGCCCTACCGAAGGGAGGGGTTTCCACCTGGGAGTCAATTAAGGTTA 2477
QY 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp----- 769
Db 2478 GTCACGCCAAGAAAATCAAGTCCNAGCTGGAGAGAAAGCGAAGACTCCATAGCT 2537
QY 770 ---SerSerValGluGlnLeuSerThrGluLeuGluProSerArgGluGluSerTyrVal 788
Db 2538 GGCTCTGGTGTAGAACATTCCTCCAGACACTGAAACCCCGTAAAGAGAATCTCTGGGTC 2597
QY 789 SerLeuLysPheLeuProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
Db 2598 TCAATCAAGAAAGTTTATCTCTGACCAAGAGAAAGGCCAGATGGGAAACAAAGACAA 2657
QY 809 AlaThrValGluAspSerGlyProValGluLeuAsnGluAspAspProAsnValProAla 828
Db 2658 GCCCTGTTGAAGACGAGCGGCGCAACAGGGGCCAACGAGATGACTCTGTATGTCGCCGCC 2717
QY 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2718 GTGGTCCCTCTGTCTGATGATGCTGTAGAAAGGAGAGAAATGGAGGCACAGCAAGCC 2777
QY 846 GlnGlyAsnThrGluLeuProGlnLeuLysGlyAlaValTyrValSerGluGluLeuSer 865
Db 2778 CAAAAAGCGCAGACGCGCCGAGAGAGGCGACCACTGAGGTGTCCAGAGAGCTCAGC 2837
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2838 GAGAGTCAGGTTTATATGATGGCAGCAGCTGTCTGACGGGACGAGGCGGTACCAT 2897
QY 886 ValGluGluArgSerProSerTyrIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2898 ATTGAAGAAAGGTCTCTCTTGGATATCTGCTTCACTGACAGAACTCTTTGAACAAGTA 2957
QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
Db 2958 GAAAGCTGAAGCCGCACTGTAACTGAGGAGGTATTGGAAAGAGAGTAATTTCAGAGAA 3017
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3018 GAACCCCGCAGGTTACTGAACCTCTGCAGAGAAAGAGAGCCCGGGCGACACGGTC 3077
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3078 GTTAGTGAGCGGAATTGACCCCGGAGGCTGTGACAGCTGCAGAAACTGCAGGCGCAT 3137
QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3138 GGTGCGGAAGAGAAACCGAAGCATCTGCTGTGAAGAGACCAACAGAAATGGTGTGAGCA 3197
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3198 GTCTCCAGTTAAACCGACTCCCGACACACACAGAGAGGCCACTCCGGTGCAGGAGGTG 3257
QY 1005 GluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaLeuLeuGlnAla 1024
Db 3258 GAAGGTGGGTACTCTGACATAGAGAGCAAGAGAGCGGACTCAAGAGGTCTCTCCAGGCA 3317
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038

```

Db 3318 GTGGCAGAAAAAGTGAAGAGGAAATCCAGCTGCTGGCAGCCGGTGGGCCAGAAAGATGTG 3377
Qy 1039 ---GlnThrValGlnArgThrGlySerIysAlaLeuGluValGluGluValGluGlu 1057
Db 3378 CTTGACGCTGTGCAGAGAGCA-----GAGGACAGAAAGACAGAGAGAG 3419
Qy 1058 AspSerGluValLeuAlaSerGluIysGluIysAspValMetProIysGlyProValGln 1077
Db 3420 CAGGCTGAAGCCCTCGGGTGTGAAGAAAGAGACGATGTGTGTGAAGTGAAGTCAAGTCTCG 3479
Qy 1078 GlnAlaGlnAlaGlnHisLeuAlaGlnIysSerGluThrGluValAlaThrProGluSer 1097
Db 3480 GAGGCAAAAACCTGAGCCTTTTACAGAGGAAAGGTGGTGGGCGACACCCCGAGAAAC 3539
Qy 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3540 TTTGAAAAAGCTCTCAAGTACAGAGACATTAAGATCCAGTGAAGCTTTAAACACTTGT 3599
Qy 1114 GlnVal-----IleIysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db 3600 CAAGCCGAAACCTTAGCTGGGTGTAATAATCAGAGATGTGATGGAACAGGCTATCCCC 3659
Qy 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3660 CCGTACTCGGTGGAACCCCTTACAGACAGTGAAGTGAAGACCCCGTAGCCGAC 3719
Qy 1148 SerAspThrAlaAspGlyThrGlnIleAspGluThrIleAspSerGlnAspSerIysAla 1167
Db 3720 TTTGACGACCCAGGACCAACCCAGAAAGCAGATGTGTGAATAATCCATGAGAGATGAG 3779
Qy 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnIysGlu 1187
Db 3780 GTGGATCTGTGTACCCAGTCAAGGGGCGACAGAGACAGAGCCGTTCTTCAAGAAAGAG 3839
Qy 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluGlnHisGlyGluGluProGly 1207
Db 3840 AGCCCTCCAGCA---CTTTCAGTTTGTGTTCAGAGAAAGAACTAAAGAAACATATAAG 3896
Qy 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProValLeu 1226
Db 3897 ATGGAAGACACTCTAGACATACAGATAAAGAGGTGTGAGTGAACCTGATTCATCTCG 3956
Qy 1227 AlaIysThrGluValGlyGlnGluGluValAspTrpLeuAspGlyGluIysValIys 1246
Db 3957 TCAAAGACTGAGGGGACTCA-----GAGGCTGACACAGTATGCTGATGAGAAACCAA 4010
Qy 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4011 GACGTACCAATTTTTCGAAGAGACTTGAAGGCTCTATGACACAGGCAATAACAGTCACTCG 4070
Qy 1260 GlnIysAlaAlaAspValThrTyraSerSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4071 GAAAAGGTACAGGAAGTTCCTTAAGGTGAAGGACAGAAAGAAAGCTGAATGTAATAAG 4130
Qy 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGluGlu 1294
Db 4131 GATGATGCTCTTGAAGCTGACAGATCGCTAACTCTCTCCATCCCGCTGAGAGAGAG 4190
Qy 1295 MetGluThrAspValGluIysGluIysAspGluThrIysProGluGlnValSerGluGlu 1314
Db 4191 ATGGTAGTTCAAGATGCAAGAGGAGAAACAGAGACAGACCAACCATGGTGAAGAG 4250
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyriGlyAspProValLeu 1333
Db 4251 AAGCTTGAGCAGAGAAACACACTGTATCCGATCTGAAGAGGTCAAGTGAAGCTCTCCAG 4310
Qy 1334 ThrLeuAspMetProSerSerGluValArgGlyValAlaLeuGlySerLeuGlySer--- 1352
Db 4311 ACAGTGAATGTCCCATCATATAGTGGGCAAGAGTGAAGTGAAGGAGAGAGCCCT 4370
Qy 1353 ProSerLeuProAspGlnAspIysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4371 CCTCCCTGCTAGGTGTAAGAGAGGAGCATATGACCAAAATTCATAGTCAAGAGCTGTGAG 4430

```

```

Qy 1373 ThrThrValThrGlnThrAlaGlnAlaValGluIysValIle-----GlnThrValVal 1390
Db 4431 GCATCATCTCACTTAACACCGCGCTGCAGAGAGAGAAAGAGCTTTAGAGAAATCCCAAC 4490
Qy 1391 IleSerGluThrGlnGluSerProGluCysValGlnAlaHisLeuLeuProAlaGluIys 1410
Db 4491 ATTTTGAACACAGGTGAACAGCTTGAAGCTTCAGAGTGCACATTTAGTTCTGGAAGAGAA 4450
Qy 1411 SerSerAlaThrGlnGluHisIleThrThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4451 TCCTGTGAAAAAAGAGAACTTTGCCCTCATCCAGGGAGAAAGATGCTGTGCCACAGAG 4610
Qy 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4611 CCCGACTGTCAAGCAAAATTCACACACAGTGAATGATCTGTACTTACCAAAAGAGCTTA 4670
Qy 1451 HisProAspLeuGlnGluIleSerAlaSerGlnArgGlnValArgSerGluGluGluIys 1470
Db 4671 AGTTCGACCTGGAAGAGAGAGAAACACATCACTGAAGTGAAGTCAAGTGAAGTCAAT 4730
Qy 1471 LysProAspAlaGlyProAspAlaAspGlyIysGluSerThrAlaIleGluValLeu 1490
Db 4731 GAGCAGGTGCTGTGCAGAGAGTC-----AAAGTGAATGTGAACATTTGAGAGAGATTTA 4784
Qy 1491 LysAlaGluProGluIleLeuGlnLeuGluSerIysSerAsnIysIleValLeuAsnVal 1510
Db 4785 GAGCCTGAAAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAACTGTCCAAAACATC 4844
Qy 1511 IleGlnThrAlaValAlaAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
Db 4845 ATCCAGACAGCCGTGACACAGTGTGTACGTACAGAAAGAACCCAGCAAAATGTTGACG 4904
Qy 1530 TyraSerSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
Db 4905 TGTGAGTTACAGACACAGAGCTCAGGTGATTAAGCTGACAGCAGAGAGCTGAGAGAA 4964
Qy 1550 TrpThrIys-----MetIysAspAla-LysMetIysHisProValProGlu 1564
Db 4965 ACGGAGAAAGAGAGAGAGAACTTCAGGCTCTGCACAGAGATGAAGAACACCAATTCTTCA 5024
Qy 1564 nProArgGlu-----AspLeuIleValLeuThrValLeuGlnAla 1577
Db 5025 GCCAAAGAGAGTCAAGTCAACCGCAGTGGAGCAAGACATTCGATATTTCCAAAGAC 5084
Qy 1577 aThrAlaGluProArgIysCysLeuProArgLeuGlnLeuIysAlaProVal 1594
Db 5085 ATGAGTGAAGCTCAAGAAAGACCATGACTGTGTGAGTGAAGAGTTCACATG 5136

RESULT 13
US-60-278-232-10143
; Sequence 10143, Application US/60278232
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0011 P
; CURRENT APPLICATION NUMBER: US/60/278, 232
; NUMBER OF SEQ ID NOS: 2001-03-30
; SOFTWARE: PERL Program
; SEQ ID NO 10143
; LENGTH: 6603
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 445786.12
US-60-278-232-10143

```


Db 4321 ACAGTGAATGCCCCATCATAGATGGGCAAGAGAGTCCAGCTTTGGAGGAGCCCT 4380
 QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
 Db 4381 CCTCCCTGCCTAGTCAAGAGGAGGAGTATGCACCAAAATTCAGTTTCAGAGCTTGAG 4440
 QY 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
 Db 4441 GCATCATCTCTTAACAGCGGCTGCAGAGAGGAGAAAGGTCTTAGAGAAATCGCCAAC 4500
 QY 1391 IleSerGlnThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
 Db 4501 ATTTAGAAACAGGTGAACGTTGAGCCCTGCAGCTGCACATTTAGTTCTTGGAGAGAAA 4560
 QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
 Db 4561 TCCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGAAGATGCTGTGCCACAGGG 4620
 QY 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
 Db 4621 CCGACTGTGAGGCAAAATCGACCAAGTATAGTATCTGCTACTACCAAGAAAGGCTTA 4680
 QY 1451 HisProAspLeuGlnGlyIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
 Db 4681 AGTTCCGACCTGGAGAGAGAGAAACCATCATCTAGTGGAGTCAATGATGATCGAT 4740
 QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
 Db 4741 GAGCAGGTGCTGCCAGGAGGTC-----AAAGTGAGTGTAGCAATTTGAGGAGATTTA 4794
 QY 1491 LysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIleValLeuAsnVal 1510
 Db 4795 GAGCCTGAAAATGGATTTTGGAACTTGAGACCAAAAGCAGTAACTTTGCCAAACATC 4854
 QY 1511 IleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
 Db 4855 ATCCAGACACCGTTGACCACTTTGACGTACAGAAAGAACAGCCACCAATTTGACG 4914
 QY 1530 TyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
 Db 4915 TCTGAGTTACAGACACAGCTCAGCTGATAAAGCTCAGACGCGGAGGAGCGTGGACAGAA 4974
 QY 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProValProGlu 1564
 Db 4975 ACGAGAGAGAGTACAGTCAACCGCAGTGGGCAACGACCATCTTGATATTTCCAAAGAC 5034
 QY 1564 nProArgGlu-----AspLeuGlnValLeuThrValLeuGluAl 1577
 Db 5035 GCCAAGAGAGTACAGTCAACCGCAGTGGGCAACGACCATCTTGATATTTCCAAAGAC 5094
 QY 1577 aTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
 Db 5095 ATGAGTGAAGCCTCAGAAAAGACCATCAGTCTTGAGGTAGAGGTTCCACTG 5146

RESULT 14
 US-60-245-228-517/c
 ; Sequence 517, Application US/60245228
 ; GENERAL INFORMATION:
 ; APPLICANT: Beasley, Ellen
 ; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL000878
 ; CURRENT APPLICATION NUMBER: US/60/245,228
 ; CURRENT FILING DATE: 2000-11-03
 ; NUMBER OF SEQ ID NOS: 630
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 517
 ; LENGTH: 6607
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; US-60-245-228-517

Alignment Scores:
 Pred. No.: 3,7e-213 Length: 6607
 Score: 4325.50 Matches: 969
 Percent Similarity: 68.83% Conservative: 188
 Best Local Similarity: 57.64% Mismatches: 408
 Query Match: 53,58% Indels: 116
 DB: 68 Gaps: 33
 US-09-902-432-4 (1-1596) x US-60-245-228-517 (1-6607)
 QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
 Db 6413 ATGGCGCGCGGAGCTCCACCGAGCAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCTCC 6354
 QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAla 39
 Db 6353 ACGCGGCTGAGCGCGAGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6303
 QY 40 GlyAspProAlaAspAla-----AspProAlaThrLysLeuProGln 53
 Db 6302 CCAGACCCCG 6243
 QY 54 LysAsnGlyGlnLeuSerValAsnGlyValAlaGluGlnGlyAspValHisValGln 73
 Db 6242 AAGAAATGGTCAGCTGTCCACCAATGGCTAGCTGAGCAAGATGAGCTCAGCCCTCCAG 6183
 QY 74 GluGluAsnGlnGluGlyGln-----80
 Db 6182 GAGGCTGACCTAAATGCCAGAAAGGAGCCCTGAACGGTCAAGGAGCCCTAAACAGCCAG 6123
 QY 81 ---GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGlu 99
 Db 6122 GAGGAAGAAGAGTCAATTTGTCACGAGGTTGGACAGAGAGACTCTGAAGATGTGAGCGAA 6063
 QY 100 LysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAsp 119
 Db 6062 AGAGACTTCGATAAAGAGATGGCTACTAAGTCAGCGGTTGTTCCACGACATCACAGATGAT 6003
 QY 120 GlyGlnGluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnValGlu 139
 Db 6002 GGCAGGAGGAGACACCCGAAATATCGAACAGATTCCTTCTTCAGAAAGCAATTTAGAA 5943
 QY 140 GluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLys 159
 Db 5942 GAGCTAACACACACCTAGTCCAGGCTAATGATATTGGATTTAAGAAGGTGTTTAAAG 5883
 QY 160 PheValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGln 179
 Db 5882 TTTGTTGGCTTTTAAATTCATCTGTGAAAAAGGATTAAGACAGAGAGAGCTGACACTGTCCAG 5823
 QY 180 LeuLeuThrValLysLysAspGluGlyGluGluValAlaGluAlaSerValGlyAlaGlyAsp 199
 Db 5822 CTACTACTGTGAAAGAAAGATGAGGGGAGGAGCA-----GCAGGGGCTGGCGAC 5772
 QY 200 HisGlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeu 219
 Db 5771 CACGAGGACCCAGCCT-----GGGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5718
 QY 220 LysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIle 239
 Db 5717 AAACAATCTACAGAGAAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5658
 QY 240 ProLeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGlu 259
 Db 5657 TCTCCCCAGCGAATCTGGCCAGCAGTG---GAGAAATGCAAGAGAGAGAGAGAGAGAG 5601
 QY 260 LysGlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnSerGlu 279
 Db 5600 AAACAAGAAAAAGAACCTAGCAAGTCTCCAGAAATCTCCGACTAGTCCCGTACCAGTGAA 5541
 QY 280 ThrThrSerSerPheLysLysPheThrHisGlyTrpAlaGlyTrpArgLysLysThr 299
 Db 5540 ACAGGATCAACCTTCAAAAAAATTTCTTCACTCAAGGTTGGCGCGCTGGCGCAAAAGACC 5481

Db 3347 CAGAGGTGAGAGTGGCTGCTGACATAGAAAGACGAGAGGCGGACTCAAGAGGTC 3288
QY 1022 LeuGlnAlaValAlaAspTysValLysGluGlnSerGlnValProAlaThr 1038
Db 3287 CTCAGGACGTGGCAGAAAAGTGAAGAGAGAAATCCAGAGTCCCTGGCACCAGGTTGGGCCA 3228
QY 1039 -----GlnThrValGlnAlaThrGlySerLysAlaLeuGluLysValGluGlu 1054
Db 3227 GAAGATGTGCTTCAGGCTGGCAGAGCA-----GAGGCGAAGA 3186
QY 1055 ValGluGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGly 1074
Db 3185 CAGAAGACAGCTGAGCGTGGGTCTGAGAAAGAGACGGATGATGATGTTGAAAGTA 3126
QY 1075 ProValGlnGluAlaGlyAlaGlnHisLeuAlaGlnGlySerGluThrGlnAlaThr 1094
Db 3125 GATGCTCAGGAGGCAAACTGAGCCTTTTACACAAGGAAGGTGGTGGGCGACACC 3066
QY 1095 ProGluSerLeuGlu-----ValProGluValThrAlaAspValAspHis-----Val 1110
Db 3065 CCAGAAGGCTTTGAAAAGCTCCTCAAGTCACAGAGAGCATAGAGTCCAGTGGCTTGTGA 3006
QY 1111 AlaThrCysGlnVal-----IleLysLeuGlnGlnLeu-----MetGluGln 1124
Db 3005 ACCACTTGTACAGCCGAACTTAGCTGGGTAAATACAGAGAGATGGTGGGAACAG 2946
QY 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
Db 2945 GCTATCCCTCCCTGACTCGGTGGAAACCCCTACAGACAGTGTGATGGAACACCCCC 2886
QY 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrLeaAspSerGlnAsp 1164
Db 2885 GTAGCGGACTTTGACGCCACCCAGGCAACACCCAGAAAGACGAGATTGTGAAATCCATGAG 2826
QY 1165 SerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluAlaAlaThrAla 1184
Db 2825 GAGAAATGAGTGCATCTGTGATCCAGTCCAGTGGGGGACAGACAGAGGAGTTCCTGCA 2766
QY 1185 GlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGlnHisGlyGlu 1204
Db 2765 CAGAAAGAGAGGCTCCAGCA---CCTTCCAGTTTGTGTGTCAGGAGAACTAAAGAA 2709
QY 1205 GluProGly---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaVal 1223
Db 2708 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTCAGTGGAAACTGTA 2649
QY 1224 ProValLeuAlaLysThrGluValGlnGlyGlnGluValAspThrPheAspGlyGlu 1243
Db 2648 TCCATTCTGTCAAGAGCTGAGGGGACTCAA-----GAGGCTACCCAGTATGCTGATGAG 2595
QY 1244 LysValLys-----GluGluGlnGluValPheValHisSerGly----- 1256
Db 2594 AAACCAAGACGCTACCATTTTTCGAGAGACTTGAGGGTCTATAGACACAGGCATAACA 2535
QY 1257 ProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGly 1276
Db 2534 GTCAGTCGGGAAAAGGTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAGAAGCTGAA 2475
QY 1277 CysGlnGluLysGluSerThrGluValGlnSer-----LeuSerLeuGlu 1291
Db 2474 TGTAAGAGGATGATGCTCTTGAACCTGCAGAGTCAGCTCAAGTCTCCTCCATCCCGGTG 2415
QY 1292 GluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnVal 1311
Db 2414 GAGAGAGATGTAGTTCAAGTCGAAAGGAGAGAAACAGACAGACCCCAACCATGTG 2355
QY 1312 SerGluGluGly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLys 1330
Db 2354 AATGAAGAGAAGCTTGAGCACGAAACAGCTGTTACCTATCTCAAGAGGTGTCAGTAAGCAG 2295
QY 1331 ProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGly 1350
Db 1350

Db 2294 CTCCTCCACACAGTGAATGTGCCCATCATAGATGGGCAAGAAAGTCAAGAGTTTGGAA 2235
QY 1351 GlySer---ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGln 1369
Db 2234 GGAAGCCCTCCTCCCTCCCTAGGTCAAGAGAGGAGGAGTATGCACCAAAAATCAAGTTTCA 2175
QY 1370 SerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----Glu 1387
Db 2174 AGCTCTGAGGCATCATCTTAACAGCGGTGCAGAGGAGGAAAAGGCTCTTAGGAGAA 2115
QY 1388 ThrValValIleSerGluThrGlyGluSerProGluCysValGlnAlaHisLeuLeuPro 1407
Db 2114 ACTGCCAACATTTAGAAACAGGTGAAACGTTGGAGCCTGCAGGTGCACATTTAGTTCTG 2055
QY 1408 AlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrVal 1427
Db 2054 GAAGAGAAATCCTCTGAAAAAATGAAGACTTTTCCGCTCATCCAGGGGAAGATGCTGTG 1995
QY 1428 ProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGlu 1447
Db 1994 CCCACAGGGCCGAGCTGTGAGGCAAAATCGACACAGTGTATGATCTCTACTACCAAG 1935
QY 1448 SerThrLeuHisProAspLeuGlnGlyIleSerAlaSerGlnArgGluArgSerGlu 1467
Db 1934 AAAGGCTTAAGTTCGACCTGGAAGAGAGAAACACATCACTGAAGTGGAGTCAGAT 1875
QY 1468 GluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGlu 1487
Db 1874 GAAGTCGATGACGAGTTCCTTCCAGGAGGTC-----AAAGTGAAGTGTAGCAATTGAG 1821
QY 1488 LysValLeuLysAlaGluProGluLysLeuGluLeuGluSerLysSerAsnLysIleVal 1507
Db 1820 GAGGATTTAGAGCCTGAAATGGGATTTTGGAACTTGGACCAAAAAGCAGTAAACTGTGTC 1761
QY 1508 LeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGlu 1526
Db 1760 CAAAACATCATCCAGACGCGTGTACACCGTGTGACAGTGTGACAGAAAGAACAGCCCGAA 1701
QY 1527 ThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluPro 1546
Db 1700 ATGTTGACGCTGAGTTACAGACACAGCTCAGCTGATATAAAGCTGACAGCCAGCGCT 1641
QY 1547 AsnArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisPr 1561
Db 1640 GGACAGGAAACGGAGAAAGAGAGAGGAAACCTCAGCGCTCTGCACAGGATGAACACCA 1581
QY 1561 OValProGlnProArgGlu-----AspLeuGlnValLeuThrVa 1574
Db 1580 ATTACTTCAGCCAAAGAGGAGTCAAGTCAACCGCAGTGGGACACGACATTTCTGATATT 1521
QY 1574 lLeuGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVa 1594
Db 1520 TCCAAAGACATAGTGAAGCCTCAGAAAGACCATGACTGTTGAGTAGAGTCCACT 1461
QY 1594 l 1594
Db 1460 G 1460

RESULT 15

US-09-644-867-6364
; Sequence 6364, Application US/09644867
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Donovan, Michael J.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Cupepper, Janice A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1173-001
; CURRENT APPLICATION NUMBER: US/09/644,867
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,061
; PRIOR FILING DATE: 1999-08-27

```

; NUMBER OF SEQ ID NOS: 8090
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6364
; LENGTH: 8660
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(8660)
; OTHER INFORMATION: n = A,T,C or G
US-09-644-867-6364

Alignment Scores:
Pred. No.: 1,56-212 Length: 8660
Score: 4316.00 Matches: 967
Percent Similarity: 68.89% Conservative: 189
Best Local Similarity: 57.63% Mismatches: 411
Query Match: 53,46# Indels: 111
DB: Gaps: 34

US-09-902-432-4 (1-1596) x US-09-644-867-6364 (1-8660)

QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
DB 183 ATGGGCGCGGAGCTCCACCGACGACCGGAGCGCGCCGACGGCGGAGGAGCTCC 242
QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
DB 243 ACCGCGCGCTAGACCCGAGCCGACGCGCGCGCGCCCTCGCGCCGAGCGCGCGACACC 302
QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLeuLeuProGluLeuAsnGly 56
DB 303 ACCGCGGACCCGCGCATCGCTGCTCGGACCCGCGCACCAAGCTCTTACAGAAAGATGGT 362
QY 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn 76
DB 363 CAGCTTCCACATCAATGCGCTGAGTGAAGATGAGTCAAGCTCAGCGAGGGGTGAC 422
QY 77 GlnGluGlyGln-----GlnGlu 82
DB 423 CTAAATGGCCAGAAAGAGCGCTGAACGCTCAAGAGCCCTTAAACAGCCAGAGAAAGAA 482
QY 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLeuAspArg 102
DB 483 GAAGTATTGTCAACGAGGTGAGACAGAGACTCTGAAGATGTGAGCGAAAGAGCTCC 542
QY 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLeuAspGlyGlnGlu 122
DB 543 GATTAAGATGAGCTACTAAGTCAAGCGGCTTTCAGACATCAAGATGATGGCGAGAG 602
QY 123 GlnThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetVal 142
DB 603 GAGAACCGGAAT---ATCGAAGCAGATTCTCTTCAAGAAAGCAATTAGAAAGCTAAAC 659
QY 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLeuLeuValPheLeuPheValGly 162
DB 660 CAACCCACTGAGTCCAGAGCTAATGATGATTGAAGAGGTGTTTAAGTTGTGGC 719
QY 163 PheLeuPheThrValIleLeuLeuAspLeuAsnGluLeuSerSerPheValGlnLeuLeuThr 182
DB 720 TTTAAATTCATCTGAAAAAGATTAAGACAGAAACCTGACACTGTCCAGCTACTC 779
QY 183 ValLeuLeuAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
DB 780 GTGAAGAAAGATGAAGGAGGAGAGCA-----GCAGGGCGCTGCGACCAACGAGAC 830
QY 203 ProSerValGluThrAlaValGlyGluSerAlaSerLeuGluSerGluLeuLeuGlnSer 222
DB 831 CCCAGCCTT-----GGGGCTGAGAAAGAGCATCCAAAGAAAGCAACCAAAACATCT 884
QY 223 ThrGluLeuGlnGluGlyThrLeuLeuGlnGlnGlnSerSerThrGluIleProLeuGln 242
DB 885 ACAGAGAAACCCGAAAGAGACCTGAAAGCTGAGCAAAAGCCAGCAAGAAATTTCTCCCCA 944

```

```

QY 243 AlaGluSerAspGlnAlaAlaGluGluGluAlaLeuAspGluGlyGluGlnGlu 262
DB 945 GCGGAATCTGGCCAAAGACTG---GAGGAATGCAAAAGAGAGAGAAACAAAGAA 1001
QY 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
DB 1002 AAAGAACTTAGAAGTCTGCAGAAATCTCCAGCTAGTCCCGTACAGAGAAACAGATCA 1061
QY 283 SerPheLeuLeuPhePheThrHisGlyTyrAlaGlyTyrPargLeuLeuSerPheLeu 302
DB 1062 ACCTTCAAAAATTTCTTCACTCAAGGTGGCGCGCTGCGCCCAAAAGACAGATTTCAG 1121
QY 303 LysSerLeuGluLeuAspAspLeuGluThrAlaGluLeuAspGlyGluGlnGluAlaGlu 322
DB 1122 AAGCCGAAGAGAGAGTGAAGTGAAGCTTCAAGAAAGAAAGAAAGAAAGCCAGAGAAAA 1181
QY 323 ValAspGluGluGluLeuGluLeuThrGluProAlaSerGluGlu----- 337
DB 1182 GTAGACAGAAAGAGAGAGAGAGAGAGAGTTCCTCCGAGAAACAGACCGCTCCGAG 1241
QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
DB 1242 CAAGCCACCCACAGAGACCGGAGAAAGTGCCACAGACCCCGGTATTCAGTTGAATAT 1301
QY 354 GlnLeuValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
DB 1302 GAGAAAGTTGAGCTGCTCCCTCAGAGAGAGCAAGTCACTGCTCGACGAGCACTTCTGA 1361
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLeuLeuMetGluAlaHisGln---Glu 392
DB 1362 AAACCTGCTCCCTTGGCCAGCAAGATGTTGATGAGAAATAGAAAGTCCACCAAGAAAG 1421
QY 393 ValValAlaGluValHisValSerThrValGluLeuThrGluGlnGlnGlyGly 412
DB 1422 GTTGTGGCCGAAGTCCAGTCCAGCCTGGAGAGAGAAACGAAAGACAG----- 1472
QY 413 GlnGluAlaGluGlyGlyValValAlaGluGlnGlyThrGlnGluSerLeuProGluGlu 432
DB 1473 -----AAACGAGGTGAGAAAGAAAGAGAGAGGTCTGCTCCAGCTGAAGAA 1517
QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluMetLysSer 452
DB 1518 TTGGTTGAAATGATGACAGAACTCAGAGAGCCGAACCTGCCAAGAGACTGTGTAAGCTC 1577
QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
DB 1578 AAAGAAAGTGTGTTTCCGAGAGAGACCTTACAGAGGAGCTGACTCAGTCTGATGAG 1637
QY 473 LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGln 492
DB 1638 AAGGTGCTGTCCAAACCCCGAGAGCGTGTGATGAGGTGGAATCTGTCACTCAACAG 1697
QY 493 GluArgIleLeuValGlnGlySerProLeuLysLeuPheSerSerSerGlyLeuLys 512
DB 1698 GAGAGAAATGAGGTGACAGGAAGTCACTTAAAGAGCTTTTACAGACACTGGCTTAAAA 1757
QY 513 LysLeuSerGlyLysLeuGlnLysGlyLysArgGlyGlyGlyAspGluGluProGly 532
DB 1758 AAGCTTTCTGAGAAAGAAACAGAAAGAGAAAG---GGAGAGAGAGACAGAGAAATCAAG 1814
QY 533 GlnTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
DB 1815 GAGCAACACTCAGAGTTCAGCCCATTTCTCCGACACCCAGAGAGCAAAAGGCGCAAGGC 1874
QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLeuGlyProLeuGlu 572
DB 1875 TCTGCTCTATCCCTTGAAGAGCCGAGAGATCACTGTCTTGAAAGAGGCTTAACCGAG 1934
QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGlyLeuLysArgGlu 592
DB 1935 GTGACAGAGAGATGGGAGAGCTGAAGAGAGAGCTACTTCCGATGAGAGAAAGAAAGAGA 1994

```

Qy	593	GlylleThrProTrpAlaSerPheIysLysMetValThrProLysLysArgValArg	612
Db	1995	GGTGTCACTCCCTGGGCATCATTTCAAAAAGATCGTGACGCCCAAGACGCTGTAGACGG	2054
Qy	613	ProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThrLeuSerSer	632
Db	2055	CCTTGGAAAGTGATTAAGAAGATGAGCTGGACAAGGTCAAGAGCGCTACCTTGTCTTCC	2114
Qy	633	ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys	652
Db	2115	ACCAGAGACACAGCCTCTGAATGCAAGAAGAAATGAAGAGGCGTGGAAAGACCAAG	2174
Qy	653	ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal	672
Db	2175	CCGGAAGAACCAAGCGCATCTGAGTGGATACCTCAGTATCTTGGAGAGCTTTAATTTGTGTG	2234
Qy	673	GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArg	692
Db	2235	GGATCATCCAAGAAAGACGAGGAGGCTCTCTTTGATGAGGAAGGGGACCAANA	2294
Qy	693	ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThr	712
Db	2295	GCAATGGGAGAGACCACCAGAAAGCTGATAGGCCGCGAAGACAAAGACGCGGACA	2354
Qy	713	AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerProGlu	732
Db	2355	GACGGATCTTCTGTTGTTCCCAAGAACATGATCAGGCGCAGGGAAGTCTCTCCCGGAG	2414
Qy	733	ProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeu	752
Db	2415	CAAGCTCGAAGCCCTACCGAAGGGAGGCGCTTCCACCTGGGAGTCAATTTAAAGGTGA	2474
Qy	753	ValThrProArgLysLysSerLysLeuGluGluLysAlaGluAsp	769
Db	2475	GTACGCCCAAGAAAATCAAGTCCAAGTCTGGAGAGAGAAAAGCAAGACTCCATAGCT	2534
Qy	770	---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal	788
Db	2535	GGGTCTGGTGTAGAACATTCACACTCCACACACTGAACCCGGTAAAGAAAGATCTCTGGGTC	2594
Qy	789	SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln	808
Db	2595	TCAATCAAGAAGTTTATCTCTGGACGAAGGAAGAAAGCCAGATCTGGAAACAAGACA	2654
Qy	809	AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla	828
Db	2655	GCCCTGTTGAGACGCGCGGCCACAGGGGCCACAGAAGATGACTCTGATGTCTCCGGCC	2714
Qy	829	ValValProLeuSerGluTyraAsnAlaValGluArgGluLysMetGlu	845
Db	2715	GTGGTCCCTCTGTCTGATGATGCTGTAGAAAGGAGMAAATGGAGGCACAGCAAGCC	2774
Qy	846	GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer	865
Db	2775	CAAAAGGCGCAGACGAGCCCGCAGAGAGGCGACCTCTGAGGTGTCTCAAGAGGACTCAGC	2834
Qy	866	LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer	885
Db	2835	GAGAGTCAGTTTCATATGATGGCGAGCTGTCTGTCAGGCGAGGGCGAGCTACCAT	2894
Qy	886	ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr	905
Db	2895	ATTGAAGAAAGGCTCTCTTCTTGATATCTGCTTCAAGTGCAGAGAACCTCTTGACAAGTA	2954
Qy	906	AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla	924
Db	2955	GAAGCTGAAGCCGCACTGTTAACTGAGGAGGTATTGGAAGAGAGAAATATTTGCAGAGAA	3014
Qy	925	GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal	944
Db	3015	GAACCCCAAGGTTACTGAACCTCTGCCAGAGAACAGAGAGGCCCGGGCGACACCGCTC	3074
Qy	945	ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu	964

Db	3075	 ::: 	GTTAGTGAGCGGAATTGACCCCGGAAGCTGTGCAGACTGCAGAAACTGCAGAGGCCATTG	3134
Qy	965	 ::: 	ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla	984
Db	3135	 ::: 	GGTCCGAAGAGAACCGAGCATCTGCTGCTGAAGAGACACACAGAAATGGTGTACAGCA	3194
Qy	985	 ::: 	ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal	1004
Db	3195	 ::: 	GTCTCCCAAGTTAACCGACTCCCGACACACACAGAGGAGGCCACTCCGGTGCAGAGGTG	3254
Qy	1005	 ::: 	GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla	1024
Db	3255	 ::: 	GAAGGTGGCTACTCCACATAGAAGACGCAAGAGCGCGACTCAAGAGGTCTCTCCAGCA	3314
Qy	1025	 ::: 	ValAlaAspLysValLysGluGluSerGlnValProAlaThr-----	1038
Db	3315	 ::: 	GTGGCAGAAAAGTGAAGAGGATCCCGACTGCTGCACCGGTGGCCAGAGAATGTG	3374
Qy	1039	 ::: 	---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu	1057
Db	3375	 ::: 	CTTCAGCCTGTGCAGAGACA-----GAGCGAGAAAGACCAAGAAAG	3416
Qy	1058	 ::: 	AspSerGluValLeuAlaSerGlnLysGluLysAspValMetProLysGlyProValGln	1077
Db	3417	 ::: 	CAGCGCTCAAGCGTGGGTCTCGAAGAAAGACGCGATAGTGTGTGAAGTAGTAGTCTCAG	3476
Qy	1078	 ::: 	GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer	1097
Db	3477	 ::: 	GAGCGAAAACCTGAGCCTTTTACAAAGGGAAGTGTGGGGCGACACCACCCGAAAGC	3536
Qy	1098	 ::: 	LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys	1113
Db	3537	 ::: 	TTTGAAAAGCTCCTCAAGTCCACAGAGACATAGAGTCCAGTGCAGCTGTGTAACCACTGT	3596
Qy	1114	 ::: 	GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla	1127
Db	3597	 ::: 	CAAGCCGAAACCTTAGCTGGGTAAATACACAGAGATGGTGTGGAATCCATGAGGAGATCAG	3656
Qy	1128	 ::: 	ProGluSerSerGlnThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp	1147
Db	3657	 ::: 	CCTGACTCGGTGAACCCCTCAGACATGAGACTGATGAAGACACCCCGCTGACCGCAC	3716
Qy	1148	 ::: 	SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla	1167
Db	3717	 ::: 	TTTGACGCACCGGCACACCCAGAAAGACGAGATTGTGGAATCCATGAGGAGATCAG	3776
Qy	1168	 ::: 	ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu	1187
Db	3777	 ::: 	GTCGCATCTGGTACCCAGTCAGGGGCACAAAGACAGAGCGAGCTTCTCTGCACAGAAAG	3836
Qy	1188	 ::: 	GluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly	1207
Db	3837	 ::: 	AGGCCTCCAGCA--CCTTCCAGTTTGTGTTCCAGGAAGAAACTAAAGAAATCAAAAG	3893
Qy	1208	 ::: 	---ArgAspValLeuGluProThrGlnGlnLeuThrAlaAlaAlaValProValLeu	1226
Db	3894	 ::: 	ATGAAGACACTCTAGAGCATACAGATAAAGAGGTGTGAGTGAAGAACTGTATCCATCTG	3953
Qy	1227	 ::: 	AlaLysThrGluValGlyGlnGluGlyGluValAspTrrLeuAspGlyGluLysValLys	1246
Db	3954	 ::: 	TCAAAGACTGAGGGACTCAA-----GAGGCTGACCATGTCTGTATGAGAAAAACCAA	4007
Qy	1247	 ::: 	-----GluGluGlnGluValPheValHisSerGly-----ProAsnSer	1259
Db	4008	 ::: 	GACTACCATTTTTTCAAGGACTTCGAGGGGTCTATAGACACAGGCATTAACAGTCGG	4067
Qy	1260	 ::: 	GlnLysAlaAlaAspValThrTyraSerSerGluValMetGlyValAlaGlyCysGlnGlu	1279
Db	4068	 ::: 	GAAGAAGTCACTGAAGTTGCCCTTTAAGGTGAAGGCACAGAAGACCTGAATGTATAAAG	4127
Qy	1280	 ::: 	LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGluGlu	1294

Db 4128 GATGATGCTTGAACGTGACAGTCAAGTCTTCATCCCGGAGAGAGAG 4187
QY 1235 MetGluThrAspValGluLysArgGluThrLysProGluIleValSerGluGlu 1314
Db 4188 ATGATGATTCAGAGTCGAAAGGAGGAAACAGAGCAGAGCAACCATGTGATGAGAG 4247
QY 1315 Gly---GluGluGluThrAlaIleProGluHisGluGlyThrTyrGlyLysProValLeu 1333
Db 4248 AAGCTTGACACGAAACAGCTGTACCTGATCTGAAAGAGTCACTAAGAGCTCTCCAG 4307
QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlySer--- 1352
Db 4308 ACAGTCAATGTCCTCCATCATAGATGGGAGGAAAGAGTCAAGCATTTGGAGAGAAAGCCT 4367
QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4368 CCTCCCTGCTGAGTCAAGAGAGGAGGAGGATGACCAAAATTCAGTTCAGAGCTCTGAG 4427
QY 1373 ThrThrValThrGlnThrAlaGluIleValGluLysValIle-----GluThrValVal 1390
Db 4428 GCATCATTCACCTTACACCGCTGACAGAGAGGAAAGGCTTACAGAGAACTGCCAAC 4487
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4488 ATTTTGAAGAACAGGTGAAACGTTGGAGCCTGCGAGCTGCACATTAGTTCTGGAGAGAAA 4547
QY 1411 SerSerAlaThrGlyGlyHisIleThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4548 TCCTCTGAAAGAAATGAGACCTTGCCTCATCCAGGGAGAGATGCTGTGCCACAGAG 4607
QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4608 CCCGACTGTACAGCAAAATCGACACAGATGATCTGCTACTACCAAGAAAGGCTTA 4667
QY 1451 HisProAspLeuGlnGlyLysIleSerAlaSerGlnArgGluArgSerGluGluAsp 1470
Db 4668 AGTTCGACCTGAGAGAGAGAGAAACCATCATCTGAAGTGAAGTCAGATGATCGAT 4727
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4728 GAGCAGGTTGCTTGCAGAGAGGTC-----AAGGTGAGTGTAGCAATTGAGGAGATTTA 4781
QY 1491 LysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIleValIleAsnVal 1510
Db 4782 GAGCCTGAAATGGAGATTGTGAACTTGAGACCAAAAGCAGTAACTTGTCCAAAACATC 4841
QY 1511 IleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
Db 4842 ATCCAGACAGCCGTTGACAGTTTGTACGTACAGAGAAACAGCCACGAAATGTTGACG 4901
QY 1530 TyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
Db 4902 TCTGAGTTACAGACACAGCTCACGATTAAGTCAAGCAGCAGCAGCGGTGAGACAGAA 4961
QY 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProValProGlu 1564
Db 4962 ACGAGAAAGAGAGAGAGACCTCAGCGCTCTGACAGATGAAACACCAATTACTTCA 5021
QY 1564 nProAArgGlu-----AspLeuGlnValLeuThrValLeuGluAla 1577
Db 5022 GCCAAAGAGAGAGTCAAGTCAACCGCAGTGGACAAAGCACATTCTGATATTTCCAAAGAC 5081
QY 1577 artpalaglnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5082 ATGAGTGAAGCTCAGAAAGACCATGACTGTGAGGTAGAAAGTTCCACTG 5133

Search completed: December 13, 2002, 06:06:25
Job time : 6800 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 00:45:36 ; Search time 1677 Seconds
(without alignments)
692.852 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 WGAGSSTQSRSEQAGSDT.....ANAPRKCLPRLQLKAPVSK 1596

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO spool/US0902432/runat_12122002_141453_25888/app query.fasta_1.1735
-DB=Pending Patents NA_New -QFMT=fastap -SUFFIX=xrpn -MINMATCH=0 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0902432 @CGN 1 1 156 @runat_12122002_141453_25888 -NCPU=6 -ICPU=3
-NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending Patents NA_New :
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8030	99.5	5236	6 US-10-152-319A-2063	Sequence 2063, Ap
2	4341.5	53.8	5931	5 US-09-724-676-12735	Sequence 12735, A
3	4341.5	53.8	5931	5 US-09-724-676A-12735	Sequence 12735, A
4	4300.5	53.3	6608	1 PCT-US02-04915-43	Sequence 43, Appl
5	4104.5	50.8	5773	5 US-09-724-676-12736	Sequence 12736, A
6	4104.5	50.8	5773	5 US-09-724-676A-12736	Sequence 12736, A
7	492	6.1	9161	6 US-10-133-937-37	Sequence 37, Appl
8	431.5	5.3	7997	6 US-10-203-138A-6112	Sequence 6112, Ap
9	427	5.3	15231	6 US-10-152-319A-1780	Sequence 1780, Ap
10	418	5.2	14335	5 US-09-724-676-25032	Sequence 25032, A
11	418	5.2	14335	5 US-09-724-676A-25032	Sequence 25032, A

12	415	5.1	8161	5 US-09-724-676-7257	Sequence 7257, Ap
13	415	5.1	8161	5 US-09-724-676A-7257	Sequence 7257, Ap
14	415	5.1	8189	5 US-09-724-676-7240	Sequence 7240, Ap
15	415	5.1	8189	5 US-09-724-676A-7240	Sequence 7240, Ap
16	415	5.1	9386	5 US-09-724-676-7256	Sequence 7256, Ap
17	415	5.1	9386	5 US-09-724-676A-7256	Sequence 7256, Ap
18	415	5.1	9414	5 US-09-724-676-7239	Sequence 7239, Ap
19	415	5.1	9414	5 US-09-724-676A-7239	Sequence 7239, Ap
20	415	5.1	9466	5 US-09-724-676-7254	Sequence 7254, Ap
21	415	5.1	9466	5 US-09-724-676A-7254	Sequence 7254, Ap
22	415	5.1	9494	5 US-09-724-676-7237	Sequence 7237, Ap
23	415	5.1	9494	5 US-09-724-676A-7237	Sequence 7237, Ap
24	415	5.1	10031	5 US-09-724-676-7255	Sequence 7255, Ap
25	415	5.1	10031	5 US-09-724-676A-7255	Sequence 7255, Ap
26	415	5.1	10059	5 US-09-724-676-7238	Sequence 7238, Ap
27	415	5.1	10059	5 US-09-724-676A-7238	Sequence 7238, Ap
28	415	5.1	10111	5 US-09-724-676-7253	Sequence 7253, Ap
29	415	5.1	10111	5 US-09-724-676A-7253	Sequence 7253, Ap
30	415	5.1	10139	5 US-09-724-676-7236	Sequence 7236, Ap
31	415	5.1	10139	5 US-09-724-676A-7236	Sequence 7236, Ap
32	415	5.1	10356	5 US-09-724-676-7252	Sequence 7252, Ap
33	415	5.1	10356	5 US-09-724-676A-7252	Sequence 7252, Ap
34	415	5.1	10384	5 US-09-724-676-7235	Sequence 7235, Ap
35	415	5.1	10384	5 US-09-724-676A-7235	Sequence 7235, Ap
36	413	5.1	5293	5 US-09-724-676-7107	Sequence 7107, Ap
37	413	5.1	5293	5 US-09-724-676A-7107	Sequence 7107, Ap
38	413	5.1	5626	5 US-09-724-676-7083	Sequence 7083, Ap
39	413	5.1	5626	5 US-09-724-676A-7083	Sequence 7083, Ap
40	412	5.1	5329	5 US-09-724-676-7104	Sequence 7104, Ap
41	412	5.1	5329	5 US-09-724-676A-7104	Sequence 7104, Ap
42	412	5.1	5662	5 US-09-724-676-7054	Sequence 7054, Ap
43	412	5.1	5662	5 US-09-724-676A-7054	Sequence 7054, Ap
44	409	5.1	5329	5 US-09-724-676-7110	Sequence 7110, Ap
45	409	5.1	5329	5 US-09-724-676A-7110	Sequence 7110, Ap

ALIGNMENTS

RESULT 1
US-10-152-319A-2063
; Sequence 2063, Application US/10152319A
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.

Db 1893 GAGCTGGAGAGGTCAGAGCGCCACCTTGTCTCTCACTGATAGCACAGTGTGAGAAATG 1952
QY 641 GluAspGluValIysThrValGlyGluGluGlnLysProGluGluProLysArgArgVal 660
Db 1953 CAAGATCAAGTCAAACTGTTGGTAGGAAACAAAGCCAGAGAAACCAAAGCGTAGGGTG 2012
QY 661 AspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArgAlaArg 680
Db 2013 GATACCTTCAGTGTCTTGGGAAGCACTGATTGTGTGGATCATCCCAAGAGAGCAAGG 2072
QY 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
Db 2073 AAGGCATCTCTTCAGATGATGAAGAGGCGCAAGGACACTGGGAGGGACAGTCACAGA 2132
QY 701 AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
Db 2133 GCAGAGAGGCGCCAGCAAGAACAAAGAACCGGAACAGACGCTGTCTCTGCAGCACCCAG 2192
QY 721 GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly 740
Db 2193 GAGCAGACCCAGCGCAAGGAAGTTCCTACCCGAGCGCGGAAGCCCTTCGGAAGGG 2252
QY 741 GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
Db 2253 GAAGGTGTCTCCACTTGGGAGTCATTTAAAGATTAGTCTACTCCAGAAAAAATCCAG 2312
QY 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
Db 2313 TCAAACTCGAAGAGAAAGCCGAAGACTCTAGTGTAGACAGTTGTCCACTGAGATCGAA 2372
QY 781 ProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArgLysLys 800
Db 2373 CCAGGTAGAGAAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACCGCGGAAGAAA 2432
QY 801 ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsn 820
Db 2433 AGGCAGATGGGAAGCAAGAACCAAGCCACTGTGGAGACTCAGGCCAGCTGGAGATAAAT 2492
QY 821 GluAspAspProAsnValProAlaValProLeuSerGluTyrAsnAlaValGluArg 840
Db 2493 GAGGACGACCCCTGATGCCAGCGCTCGTCTGTCTGATGATGATGATGATGATGATGATG 2552
QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal 860
Db 2553 GAGAAGATGAAGCCCAAGGGGAATCGCGAGCTGCCCGAGCTGCTGGGGGCTGTGTACGTG 2612
QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
Db 2613 TCCGAGGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTCCGAGTCAATTGATGGACC 2672
QY 881 ArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrGlu 900
Db 2673 AGGCAGTCAACAGTGTGAAGAGCGGTCTCTTCGTGGATATCCGCTTCGGTAACAGAA 2732
QY 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
Db 2733 CTTCTTTGAACACACAGCGGGAGAGCCATGCCACTGTTGAAGAGGTCACTGAAAAAGAC 2792
QY 921 IleIleAlaGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
Db 2793 ATCATTCAGAAGAAATCTCTGTCTCACCAGACGTTTACAGAGGTTAAGATGCCCAT 2852
QY 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
Db 2853 GACGACATGTTCACCACTGAAGTGGATTTTCACTCAGAACTGTGACAGGCCACAGAGACC 2912
QY 961 SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrAsp 980
Db 2913 TCAGAGGCTCTCCGTACTGAAGAAGTTACCGAAGCATCGGGGGCCGAAGAGACACACAGAC 2972
QY 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
Db 2973 ATGTTGTCCGAGTTTCCCACTGACTGACTTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAG 3032

QY 1001 ValGlnGluValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAla 1020
Db 3033 GTTCAGAGCTAGAGAGTGTGTGTGTAGATACAGAAAGAGAGAGCGCCAGACGAGCC 3092
QY 1021 IleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThr 1040
Db 3093 ATCCTCCAGCCGTTGCAGACAAGGTGAAGAGAGTCCCAAGTGTCTGCAACCCAGACT 3152
QY 1041 ValGlnArgThrGlySerLysAlaLeuGluLysValGluValGluGluAspSerGlu 1060
Db 3153 GTCACAGAAACCGGGTCNAAGGACCTGGAGAGGTTGAGAGGTTAGAGAGGACTCCGAA 3212
QY 1061 ValLeuAlaSerGlnLysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
Db 3213 GTCTGTGCTTCGAGAGAAAGAGAGGACGTTATGCCGAAAGGACCCGTCGAGGAAGCTGA 3272
QY 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
Db 3273 GCTGAGCATCTTCACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCTTTGAAGTT 3332
QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
Db 3333 CCTGAAGTCACGGCAGATGTAGACCATGTGCCACCGTCCAGGTTATCAAGCTCCAGCAG 3392
QY 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
Db 3393 CTGATGGAAACAGCCGCTGGCCCTGAGTGCATCCGAAACCTTGACACAGCTGAGACAAAT 3452
QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
Db 3453 GGAAGCAGCTCCCTTAGCAGATTCCAGACTGCAGATGGGACACAGCAAGATGAACCATTT 3512
QY 1161 AspSerGlnAspSerLysAlaThrAlaValArgGlnSerGlnValThrGluGluGlu 1180
Db 3513 GACAGCCAGGACAGTAAGCCACTGCAGCTGTCCAGCAGTCCAGGTCACAGAAAGAGAG 3572
QY 1181 AlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGlu 1200
Db 3573 CGCGCTACTGCTCAGAAAGAGGAGCCTTCGACACTACTTAATAATGTTCCAGCCAGGAA 3632
QY 1201 GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAla 1220
Db 3633 GAACATGGGGGAAGAACAGAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTACTGCT 3692
QY 1221 AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyValAspTrpLeu 1240
Db 3693 GCAGCCGTGCCCTTCTGGCAAGACTGAGTGGGTCAAGAGGGTGAGGTTGACTGGTTG 3752
QY 1241 AspGlyGluLysValLysGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
Db 3753 GATGGAGAAAAAGTCAAAAGAAACAGGAGGTGTTGTACACTCTGGACCCCAACAGTCAA 3812
QY 1261 LysAlaAlaAspValThrTyrAspSerGluValMetClyValAlaGlyCysGlnGluLys 1280
Db 3813 AAGGCTGCTGATGTGACATATGACAGTGAAGTGTGGAGTGGCGGGTGTCTGAGGAAAG 3872
QY 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluGlyValMetGluThrAspValGlu 1300
Db 3873 GAGGTACTGAAGTGCAGAGTCTTAGCTGGAGAGGAGAGATGGAAATCTGACGTTGAA 3932
QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGlyGluGlnGluThrAla 1320
Db 3933 AAGGAGAAAAAGGAGACAAAGCCAGAGCAAGTGAAGTGAAGAGTGAAGAGGAAACAGCC 3992
QY 1321 AlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSer 1340
Db 3993 GCTCCTGAGCATCAAGGAACCTACGGGAAGCCAGTCTCTGACATTTGACATGCCAGCTCA 4052
QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyLysSerProSerLeuProAspGlnAspLys 1360
Db 4053 GAGAGGGGGAAGGCACTGGGAAGCCTTGGGAAGGAGCCCTTCTCTCCAGAGCAAGCAAA 4112

Qy 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGln 1380
Db 4113 GCAGGTGCAATGAGAGTTCAAGTTCACAACTGAGACACAGCTCATCAACACGACGAA 4172
Qy 1381 AlaValGluValValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
Db 4173 GCTGGGAAAAGTCATGAAACGGTTGATTTCAGACACAGGTGAAGTCAAGAGTGT 4232
Qy 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisThrThrLeu 1420
Db 4233 GTAGGTGACACTTATTCACAGCTGGAAGTCTCTGCAACGGGTGGCCACAGCTGCTTT 4292
Qy 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGlnSerIleProIle 1440
Db 4293 CAGCATGCGAGAGACACGGTACCCCTGGGGCTGAGTCTCAGGCAAGATTCATCCCAATC 4352
Qy 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAla 1460
Db 4353 ATAGTACTCCTGCTCTCTTAAAGCACCCCTACATCTCTGACCTTCAAGGAAATTAAGCGCA 4412
Qy 1461 SerGlnArgGluArgSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
Db 4413 TCCACAGAGAGACGATCAGAGAGAGACCAAGCCAGATGCTGTCCTGATGCTGACGCGC 4472
Qy 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGlnIleLeuGluLeuGlu 1500
Db 4473 AAGGAGAGTACAGCAATCAAAAAGCTCTCAAGGCTGAACCTGAGTCTTGAACTTGAG 4532
Qy 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
Db 4533 AGTAAGAGCAACAAGATTGCTGTAACGTCATTCAGACAGCCGTTGACCAAGTTCGACGT 4592
Qy 1521 ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysArg 1540
Db 4593 ACAGAAACAGCCCGGAAACTCATGCTTATGATTTCACAGACCAGGTTCCTGCAATGGCAG 4652
Qy 1541 LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
Db 4653 GCTGACAGAGGAGCCCAACAGATCTGACAAAAATTAAGTTCCCAAGATGAAACAC 4712
Qy 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGlnValLeuAlaTrpAlaGln 1580
Db 4713 CCAGTCCGCCACCCACAGAGAGACTTGCAAGTCTGACCGTCTTGAGAGGCAATGGGCTCAG 4772
Qy 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 4773 CCTCGAAATGCTTGCGCGCTTGCAAGTGAAGCGCCGCTGTCAAG 4820

RESULT 2
US-09-724-676-12735
; Sequence 12735, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12735
; LENGTH: 5931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12735

Alignment Scores:
Pred. No.: 3,65e-172 Length: 5931
Score: 4341.50 Matches: 970
Percent Similarity: 69.13% Conservatave: 190
Best Local Similarity: 57.81% Mismatches: 408
Query Match: 53.78% Indels: 110
Gaps: 33

US-09-902-432-4 (1-1596) x US-09-724-676-12735 (1-5931)
Qy 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
Db 192 ATGGCGCGCGGAGAGTCCACCGAGAGCGCAGCCCGGAGCAGCCCGAGGGGAGCTCC 251
Qy 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
Db 252 AGCGCGGCTGAGCCCGAGCCCGAGCCGCGCGGCGCGCTTGCGCGAGGCGCGCCAGACACC 311
Qy 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
Db 312 ACCCGGAGCCCGGACATGCTGCTGCGACCCCGCACCAAGGCTCTTACAGAAAGATGT 371
Qy 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGln 76
Db 372 CAGCTGTCCACATCAATGAGGCTGAGTGAAGATGAGTCAAGCTTCAGAGGCGGTGAC 431
Qy 77 GlnGluGlnGln-----GlnGlu 82
Db 432 CTAAATGGCCAGAAAGAGCCCTGAAACGCTCAAGAGCCCTTAACACGCGAGAGAGAA 491
Qy 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
Db 492 GAAGTCATTGTCACAGAGGTTGACAGAGAGACTTGAAGATGTGAGGGAAGAGACTCC 551
Qy 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu 122
Db 552 GATAAAGAGATGGCTTACTAAGTACGCGGTGTTCCAGCATCACAGATGATGGCAGAG 611
Qy 123 GlnThrSerGluIleIleGlnGlnIleProAlaSerGluAsnAsnValGlnGluMetAla 142
Db 612 GAGACACCCGAAATATGAAACAGATTCCTTCTTGAAGACCAATTTTGAAGAGCTAAC 671
Qy 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
Db 672 CAACCCACGTAGTCCACGAGCTTATGATTTGAATTAAGAGGTGTTAAAGTTCTTGGC 731
Qy 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
Db 732 TTTAAATTCACTGTGAAAAAGATTAAGACAGAGAACTGACACTGTCCAGTACTACT 791
Qy 183 ValLysLysAspGluGluGluValAlaGluLysAspValGlyAlaGlyAspHisGlnGlu 202
Db 792 GTAAAGAAAGATGAAGGGGAGGAGCA-----GAGGGGCTGGGACCCACAGGAC 842
Qy 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
Db 843 CCCAGCTT-----GGGGCTGGAGAGACAGCTTCCAAAGAAAGGAAACCCAAACATCT 896
Qy 223 ThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIleProLeuGln 242
Db 897 ACAGAGAAACCCAGAGAGACCTTGAAAGCTGAGCAAAACGACGAGAAATTTCTCCCCCA 956
Qy 243 AlaGluSerAspGlnAlaAlaGluGluGluLysAspGluGluGluLysGlnGlu 262
Db 957 GCGGAATCTGGCCAGGACAGTG--GAGGAATGCCAAAGGAGGAGGAGAAACAAAGAA 1013
Qy 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
Db 1014 AAAGAACTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTGCACAGTGAAGACAGATCA 1073
Qy 283 SerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSerPheLys 302
Db 1074 ACCTTCAAAAAATTTCTTCACTCAAGTTGGGCGGCTGGCCGCAAAAAGACAGATTTCAGG 1133
Qy 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
Db 1134 AAGCCGAAAGGAGAGGAGTGAAGCTTCAAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1193
Qy 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
Db 1194 GTAGACACAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1253

QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
Db 1254 CAAGCCACCACAGGAGCGCGAGAAAGTGCACAGCCCGGTTATCAGCTGAATAT 1313
QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1314 GAGAAAGTGAAGTGCCTCAGAGGAGCAAGTCAAGTGGCTCGCGGACCTTCTGAAGAG 1373
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
Db 1374 AAACCTGCTCCGTTGGCGACAGAAAGTGTGATGAGAAAATAGAAGTCCACCAAGAGAG 1433
QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGlyGly 412
Db 1434 GTTGTGGCCGAAGTCCAGCTCAGCACCGTGGAGAGAGAACCCGAAGAGCAG----- 1484
QY 413 GlyGluAlaGluGlyGlyValValValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1485 -----AAAACGGAGGTGGAAGAAACACAGCGGTCTGTGCCAGCTGAAGAA 1529
QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSer 452
Db 1530 TTGTTGAAATGGATGCAGAAACCTCAGAAAGCCGAACCTGCCAAGGAGCTGGTGAAGCTC 1589
QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1590 AAAGAAACGTGTGTTCCGAGAGGAGCCCTACACAGGAGCTACCTCAGTCTCTGATGAG 1649
QY 473 LysThrLeuProLysHisProGluGlyLeValSerGluValGluMetLeuSerSerGln 492
Db 1650 AAGTGTCTGCCAAACCCCGAAGGGGTGTGTAGTGAGGTGGAATGCTGTCAATCAG 1709
QY 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerGlyLeuLys 512
Db 1710 GAGAAATGAAGTGCAGGAGAGTCCACTAAGAGAGCTTTTACCAGACTGGCTTAAAA 1769
QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGly 532
Db 1770 AAGCTTTCTGGAAGAAACAGAAAGGGAAGA---GGAGGAGGAGACGGAATCAGGG 1826
QY 533 GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1827 GAGCACTCAGGTTCAGCGCGATTCTCCGACAGCCAGGAGGAGCAAAAGGGCGAGGC 1886
QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1887 TCTGCCTCATCCCTGAGGAGCCGAGAGATCAGCTGTCTGGAAAGGGCTTAGCCGAG 1946
QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1947 GTGCAGCAGGATGGGAAGCTGAAGAAGGAGTACTTCCGATGGAGAGAAAAGAGAA 2006
QY 593 GlyIleThrProTrpAlaSerPheLysMetValThrProLysLysArgValArgArg 612
Db 2007 GGTGTCACTCCCTGGGCATCATTTCAAAGAATGTGTACGCCCAAGACGCTGTAGACGG 2066
QY 613 ProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
Db 2067 CCTTCGGAAGTGAATAAGAGATGAGCTGCAGAGTCAAGAGCGCTACTTGTCTTCC 2126
QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGlnLys 652
Db 2127 ACCGAGACACAGCTCTGAAATGCAAGAAAGAAATGAAGGAGCGGTGGAAGGCCAAAG 2186
QY 653 ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
Db 2187 CCGAAGAACCAAGCCAGGTGGATACCTCAGTATCTTGGGAAGCTTTAAATTTGTGTG 2246
QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg 692
Db 2247 GGATCATCCAAAGAAAGCAAGGAGGAGGTCTCTTCTGATGGAAGGAGGACCAAA 2306

QY 593 ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2307 GCAATGGGAGGAGACACCAGAAAGCTGATGAGCCGGAAGAACAAAGACAGCGGACA 2366
QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerProGlu 732
Db 2367 GACGGATCTCTGCTGGTCCCAAGAACATGATCCAGGGCAGGAAGTCTCTCCCGGAG 2426
QY 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeu 752
Db 2427 CAAGCTGGAAGCCCTACCGAAGGGAGGGCTTTCCACCTGGGAGTCATTTAAAGGTTA 2486
QY 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp----- 769
Db 2487 GTCACGCCCAAGAAAAAATCAAAAGTCCAAGCTGGAAGAGAAAAAGAAAGACTCCATAGCT 2546
QY 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal 788
Db 2547 GGTCTGTGTAGAACATTCCACTCCAGACACTGAACCCCGTAAAGAGNANTCTGGGTC 2606
QY 789 SerIleLysPheIleProGlyArgArgLysArgAlaAspGlyLysGlnGln 808
Db 2607 TCAATCAAGAAGTATTCTCTGCACGAAGGAAGAAAGGCCAGATGGGAAACAAGAACAA 2666
QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2667 GCCCTCTGTTGAAGACACGCGGCCAACAGGGGCCAACGAAGATGACTCTGTGTGCCGCC 2726
QY 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2727 GTGTCCTCTGTCTGAGTATGATCTGTAGAAAGGAGAAAAATGGAGGCACAGCAAGGCC 2786
QY 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2787 CAABAAGCCGACAGCAGCCGAGCAGAGGAGCCACTGAGGTGTCAGAGGAGCTCAGC 2846
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2847 GAGAGTCAGGTTCATATGATGGCAGCAGCTGCTGACGGGACGAGGCGAGCTACCAT 2906
QY 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2907 ATTGAAGAAGGTCTCCTCTTGATATCTGCTTCAAGTACAGAACCTCTTGAACAAGTA 2966
QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
Db 2967 GAAGCTGAAGCCGACACTGTTAACTGAGGAGGTATTGGAAGAGAGTAATTCAGAGAA 3026
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3027 GAACCCCCACCGTTACTGAACCTCTGCCAGAGAACAGAGAGGCCCGGGCGCACCGGTC 3086
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3087 GTTAGTGAGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGCCCAT 3146
QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3147 GGTGCCGAGAGAGAACCGAGCATCTGCTGCTGAAGAGACCAAGAAATGGTGTACGA 3206
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3207 GTCTCCAGTTAACCGACTCCCCAGACACCCAGAGAGGCCACTCCGGTGCAGGAGGTG 3266
QY 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3267 GAAAGTGGCGTACTCAGCATAGAGAGCAAGAGAGCGGACTCAAGAGGTCTCTCCAGGCA 3326
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3327 GTGGCAGAAAAGTGAAGAGGAATCCAGCTGCTCCCTGGCACCGGTGGCCAGAGATGTG 3386
QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057

```

Db 3387 CTTCAGCTGTGACAGAGCA-----GAGCGAGAAAGACCAGAGAG 3428
Qy 1058 Asperglutinalleualaserglutylsaspvalmetprolysglyprovalgin 1077
Db 3429 CAGCGTGAAGCGTGGGTCGTGAAGAAAGACGAGATGTGTGAAAGTAGTGCTGAG 3488
Qy 1078 Glu1aaglyalagluhi1leualasergluthtgylvalalathrprogluser 1097
Db 3489 GAGCGAAAAACTGACCTTTTACACAGAGGAGGTGTGGGACAGCCAGGAAAC 3548
Qy 1098 Leuglu-----ValProgluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3549 TTTGAAAAAGCTCTCAATGTCACAGAGCATAGACTCCAGTGAGTTTAACTTGT 3608
Qy 1114 GluVal-----Llelyseuglnglnleu-----Metglu1a1aValAla 1127
Db 3609 CAGCGCAAACTTACGTGGGTTAAATGACAGAGATGTGTGAAACAGGCTATCCCC 3668
Qy 1128 ProgluserSerGluThrleuthrAspSerGluThrAsnGlySerThrProleuAlaasp 1147
Db 3669 CCGTACTCGGTGAAACCCCTACAGACGTGAGTGAAGCAACCCCGTAGCCGAC 3728
Qy 1148 SerAspThrAlaaspGlyThrGlnGlnAspGlyThrIleAspSerGlnAspSerLysAla 1167
Db 3729 TTTGACGACCCAGGACCAACCAAGAAAGCAAGATTGTGAATCCATGAGAAATGAG 3788
Qy 1168 ThrAlaAlaValArgLuserGlnValThrGluGluValAlaAlaThrAlaGlnysGlu 1187
Db 3789 CTCGATCTCGTACCAAGTCAGAGGGGACAGAAAGCAAGAGCTTCTGACAGAAAGAG 3848
Qy 1188 GluProSerThrleuProAsnAsnValProAlaGlnGlnGlnHisGlyGluGluProGly 1207
Db 3849 AGGCTCCAGCA-----CTTCCAGTTTGTGTTCCAGGAAACCTAAGAACATCCAAAG 3905
Qy 1208 ---ArgAspValleugluProThrGlnGlnGlnGlnValAlaAlaAlaValProValleu 1226
Db 3906 ATGAGAGACCTCTAGCATACAGATTAAGAGGTCTAGTGAACCTATTCATTCTCG 3965
Qy 1227 AlaIythrGluValGlyGlnGlnGluValAspTrpLeuAspGlyGluValys 1246
Db 3966 TCAAAAGACTGAGGGAGCTCA-----GAGGCTGACCAATAGCTGATGAGAAACAA 4019
Qy 1247 -----GluGlnGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4020 GACGTACCATTTTTCAGAGACTTGGGGGTCTATGACACAGGCACTTAACGTACTCG 4079
Qy 1260 GlnIythrAlaAlaAspValThrTyraSpSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4080 GAAAAGGTCACTGAAGTTGCCCTTAAGGTGAAGGACAGAAAGAACTGAATGTAAAG 4139
Qy 1280 LysGluSerThrGluValGlnSer-----LeuSerleuglngluGluGlu 1294
Db 4140 GATGATGCTCTGTAAGTCAAGTCAAGTCTCCTCCATCCCCCGTGAGAGAGAG 4199
Qy 1295 MetGluThrAspValGlnIythrGlnIythrAspGluThrLysProGluGlnValSerGluGlu 1314
Db 4200 ATGTAGATTCAAGTCGAAAGGAGAGAAACAGAAAGCAAGCCATGATGATGAAGAG 4259
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrclyLysProValleu 1333
Db 4260 AAGCTTGACACAGAAACACTGTATCCGTATGAAAGGTCAATGAAGCTCCCTCAG 4319
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerleuglyGlySer--- 1352
Db 4320 AAGAGTGAATGCTCCATCTAATATGGGCAAGGAAGTCAAGAGTTTGAAGAAACCT 4379
Qy 1353 ProSerleuProAspGlnAspLysAlaGlyCysIleGluValGlnValInserleuAsp 1372
Db 4380 CCTCCCTGCTGAGTCAAGAGAGAGGACGATATCACCAAAATTCAGTCAAGAGCTCTGAG 4439
Qy 1373 ThrThrValThrGlnThrAlaGluAlaValGluValIle-----GluThrValVal 1390

```

```

Db 4440 GCATCATTCACCTTAACAGCGGCTGACAGAGAGGAAAAAGTCTTAAGAGAACTGCCAAC 4499
Qy 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4500 ATTTAGAAACAGAGTGAACCTTGAGCCTGACAGTGCACATTTAGTTCTGGAAGAGAA 4559
Qy 1411 SerSerAlaThrGlyGlyHisTrpThrleuGlnHisAlaGluAspThrValProleuGly 1430
Db 4560 TCCTTGAAAAAATGAAGACTTGGCCCTCATCCAGGGAGAGTGTGTGCCACAGGG 4619
Qy 1431 ProgluserGlnAlaGluSerIleProIleValThrProAlaProgluserThrleu 1450
Db 4620 CCGCATCTGACGCAAAATTCACACAGTATGATCTGTGATCTGATACCAAGAAAGCTTA 4679
Qy 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluLys 1470
Db 4680 AGTTCGACCTGGAAGGAGAGAAAAACATCATCAGTAAGTGAAGTCAATGATGAT 4739
Qy 1471 LysProAspAlaGlyProAspAlaaspGlyLysGluSerThrAlaIleGluValleu 1490
Db 4740 GAGCAGGTGCTTGCCAGAGAGTC-----AAAGTGAAGTGAACATTTGAGAGGATTTA 4793
Qy 1491 LysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIleValleuAsnVal 1510
Db 4794 GAGCTGAAATGGATTTTGGAACTTGAGACCAAAAGCAGTAACTTGTCCAAACATC 4853
Qy 1511 IleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
Db 4854 ATCCAGACAGCCGTGACACAGTTGTGATCAGAAAGAAACAGCACGGAATGTTGACG 4913
Qy 1530 TyraSpSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
Db 4914 TCTGATTACAGACCAAGCTCACTGATTAAGCTGACACCCAGAGCGTGGAGAAA 4973
Qy 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProvalProGlu 1564
Db 4974 ACGGAGAAAGAGAGAGAGAACCTCAGGCTCTGACAGATGAAACCAATTACTTCA 5033
Qy 1564 nProArgGlu-----AspleuGlnValleuThrValleuGluAl 1577
Db 5034 GCCAAAGAGAGTCAAGACTCAACCGCAGTGGAGACAGCACTTGTGATTTCCAAAGAC 5093
Qy 1577 aTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProval 1594
Db 5094 ATGACTGAAGCTCAGAAAAGCAATGACTGTGAGTGAAGGTTCCACTG 5145

RESULT 3
US-09-724-676A-12735
; Sequence 12735, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12735
; LENGTH: 5931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12735

Alignment Scores:
Pred. No.: 3,65e-172 Length: 5931
Score: 4341.50 Matches: 970
Percent Similarity: 69.13% Conservative: 190
Best Local Similarity: 57.81% Mismatches: 408
Query Match: 53.78% Indels: 110
DB: 5 Gaps: 33

```

US-09-902-432-4 (1-1596) x US-09-724-676A-12735 (1-5931)

Qy	1	MecGlyAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp	19
Db	192	ATGGCGCCGGAGCTCCACCGAGCAGCGCAGCCGGAGCAGCCCGCGAGGGAGTCC	251
Qy	20	ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla	38
Db	252	ACGCGCGCTGAGCCGAGCCACGCGCGCGGCCCTCTGGCCGAGCGCGCCGACAC	311
Qy	39	AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly	56
Db	312	ACCGCGACCCCGCCATCGCTCGCTCGGACCCGCCACCAAGCTCTCAGACGAATGGT	371
Qy	57	GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsn	76
Db	372	CAGTGTGCCACCATCAATGCGGTAGCTGAGCAAGATGAGCTCAGCGCTCCAGAGGGT	431
Qy	77	GlnGluGlyGln-----GluGlu	82
Db	432	CTAAATGGCCAGAAAGAGCCCTGAACGGTCAAGGAGCCCTAAACAGCCGAGGAGAA	491
Qy	83	GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg	102
Db	492	GAAGTCATTCTCAGAGGTTGGCAGAGAGACTCTGAAGATGTGCGCGAAAGACTCC	551
Qy	103	ValGluGluMetAlaAlaAsnSerThrAlaValGluAspPileThrLysAspGlyGlnGlu	122
Db	552	GATAAAGAGATGGCTAAAGTCAGCGGTTGTTTCAGCATCACAGATGATGGGCAGGAG	611
Qy	123	GluThrSerGluIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetVal	142
Db	612	GAGACACCCGAAATAATCGAACAGATTCCTCTTCAGAAAGCAATTTAGAAGAGCTA	671
Qy	143	GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly	162
Db	672	CNACCCACTGAGTCCAGGCTAATGATATGATTTAGAAGGTGTTTAAAGTTGTGGC	731
Qy	163	PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr	182
Db	732	TTTAAATTTCACTGTGAAAAGGATAAGACAGAGAGCCCTGCACACTGTCCAGCTACT	791
Qy	183	ValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu	202
Db	792	GTGAAGAAAGATGAAGGGAGGGAGCA-----GCAGGGGCTGGCGCACCCAGGAC	842
Qy	203	ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLysGlnSer	222
Db	843	CCAGCCTT-----GGGGCTGGAGAGCAGCATCCNAAAGAAAGCGAACCCNACATCT	896
Qy	223	ThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIleProLeuGln	242
Db	897	ACAGAGAAACCCGAAAGAGACCCCTGAAGCGTGAGCAAAAGCCACGCGAATAATTTCT	956
Qy	243	AlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGluLysGlnGlu	262
Db	957	GCCGAATCTGGCCACAGCGTG---GAGGAATGCCAAGAGGAGGAGAGAGAGAAACA	1013
Qy	263	LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer	282
Db	1014	AAAGAACTAGCAAGCTCGAGAAATCTCCGACTAGTCCCGTCCAGTCCAGTGAACAGGAT	1073
Qy	283	SerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSerPheLys	302
Db	1074	ACCTTCAAAAAATTTCTCACTCAAGGTTGGCGCGCTGGCGCAAAAGACCAAGTTTCAG	1133
Qy	303	LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys	322
Db	1134	AAGCCGAAGAGAGATGAAGTGGAAAGCTTCAGAGAAGAAAAAGGAAACAAGAGCCGA	1193
Qy	323	ValAspGluGluLysGlnLysThrGluProAlaSerGluGlu-----GluGlu	337
Db	1194	GTAGACACAGAGAGACGGAAGGCGAGGTTGGCTCCGAGAAATCTGACCGCTCCGAG	1253
Qy	338	-----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspThr	353

Db	1254	CAAGCCACCCACGAGAGCGGCAGAAAAGTGCCACGAGGCCCGGTATCAGCTGAATAT	1313
Qy	354	GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu	373
Db	1314	GAGAAAGTTGAGCTGCCCTCAGAGAGGCAAGTCACTGCTCGCAGGACCTTCTGAAGAG	1373
Qy	374	LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---	392
Db	1374	AAACCTGCTCCGTTGGCGACAGAAGTGTTTGATGAGAAAATAGAAAGTCCACCAAGAGAG	1433
Qy	393	ValValAlaGluValHisValSerThrValGluLysThrGluGluGluGlnGlyGly	412
Db	1434	GTTGTGGCCAGTCCACGCTCAGCACCGGTGGAGGAGAGAACCGAAGAGCAG-----	1484
Qy	413	GlyGluAlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProProGluLys	432
Db	1485	-----AAAACGAGGTGGAGAAACAGCAGGGTCTGTGCCAGCTGAAGAA	1529
Qy	433	LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer	452
Db	1530	TTGGTTGAATGGATGCAGAACCTCAGGAAGCCGAACCTGCCAAGGAGCTGGTGAACCTC	1589
Qy	453	ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu	472
Db	1590	AAAGAAACGTGTGTTTCCGAGAGGAGACCCTACACAGGAGGAGTCACTCAGTCTGATGAG	1649
Qy	473	LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGln	492
Db	1650	AAGTGTCTGCCAAACCCCCGAAGCGTGTGAGTGAGGTGGAAATGCTGTCTATCAG	1709
Qy	493	GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys	512
Db	1710	GAGAGATGAAGTGCAGGGAAGTCCACTAAGAAGCTTTTACCAGCACTGGCTTAA	1769
Qy	513	LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysAspGluGluProGly	532
Db	1770	AAGCTTTCTGGAAGAAACAGAAAGGGAAGA---CGAGGAGGAGACAGAGGAATCAGGG	1826
Qy	533	GluTyrlGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer	552
Db	1827	GAGCACTCAGGTCTCCAGCCGATTCTCCGACACCCAGGAGGAGCAAAAGGGCGAGGC	1886
Qy	553	SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu	572
Db	1887	TCTGCCTCATCCCTCAGGAGCCCGAGGAGATCAGTGTCTTGAAAGGGCTTAGCCGAG	1946
Qy	573	AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu	592
Db	1947	GTGCAGCAGGATGGGAAGCTGAAGAAGGAGCTACTCCGATGCAGAGAAAAAAGAGAA	2006
Qy	593	GlyIleThrProThrAlaSerPheLysLysMetValThrProLysLysArgValArgArg	612
Db	2007	GGTGCTACTCCCTGGGCATCATTTCAAAAAGATGGTGACGCCCAAGAGCGTGTAGACGG	2066
Qy	613	ProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThrLeuSerSer	632
Db	2067	CTTTCGAAAGTGATAAAGAGATGAGCTGGACAGGTCAAGAGCGCTACCTTGTCTTCC	2126
Qy	633	ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys	652
Db	2127	ACCAGAGACAGCCCTCTGAATGCAAGAAGAAATGAAAGGGAGCGCTGGAAGAGCCCAAG	2186
Qy	653	ProGluLupProLysArgValAspThrSerValSerThrGluAlaLeuIleCysVal	672
Db	2187	CCGGAAGAACCAAGCGAAGTGGATACCTCAGTATCTTGGGAAGCTTTAATTTGTGTG	2246
Qy	673	GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg	692
Db	2247	GGATCATCCAAAGAAACAGCAGAGAGGGTCTCTTCTGATGAGAAAGGGGGHACAAAA	2306
Qy	693	ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThr	712

Dh 2307 GCATGGAGGAGACACACAGAAAGCTGATGAGCCCGGAAAAAGACAAGAGAGCGGAGACA 2366
Qy 713 AspAlaValProAlaSerThrGlnGlnGlnAspGlnAlaGlnGlySerSerProGln 732
Dh 2367 GACGGGATCTCTGCTGCTCCCAAGAACATGATCCAGGCGAGGAAAGTCTCTCCCGGAG 2426
Qy 733 ProAlaGlySerProSerGlnGlnGlnValSerThrThrGlnSerPheLysArgLeu 752
Dh 2427 CAAGCTGGAAAGCTTACCAAGAGGAGGCGTTTCCACTGGGAGTCATTAAAGGTTA 2486
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGlnGlnLysAlaGlnAsp 769
Dh 2487 GTCAAGCCCAAGAAAAAATCAAGTCCAGCTGGAAGAAAAACCAAGACCTCCATAGCT 2546
Qy 770 ---SerSerValGlnGlnLeuSerThrGlnGlnProSerArgGlnGlnSerThrVal 788
Dh 2547 GGGCTGGGTGAGAAACATTCACCTCCAGACACTGAACCCGGTAAAGAAATCTGGGCTC 2606
Qy 789 SerLysLysPheLeuProGlnArgArgLysLysArgAlaAspGlnLysGlnGln 808
Dh 2607 TCAATCAAGAAAGTTATATCTTGAGAGAGAAAGAAAGCCAGATGGAAACAAGAACAA 2666
Qy 809 AlaThrValGlnAspSerGlnProValGlnLysGlnLysAspAspProAsnValProAla 828
Dh 2667 GCCCTGTTGAAGACCGCAGGCGCAACAGGGGCCAAGAAAGTACCTGATGCTCCGGCC 2726
Qy 829 ValValProLeuSerGlnLysAsnAlaValGlnArgGlnLysMetGln 845
Dh 2727 GTGGTCCCTCTGTGATGATGATGCTGAGAAAGGAGAAATGAGGACACAGCAAGCC 2786
Qy 846 GlnGlnAsnThrGlnLeuProGlnLeuGlnValValThrValSerGlnLeuLeuSer 865
Dh 2787 CAAAAAGCGCAGAGAGCGCCAGAGCAGAAAGCAGCACTGAGGTGTCAGAGGCTCAGC 2846
Qy 866 LysThrLeuValHisThrAlaSerValAlaValLysAspGlnThrArgAlaValThrSer 885
Dh 2847 GAGAGTCAGGTTCAATGATGATGAGCAGCTGCTGAGAGGAGCAGAGGCGAGTTCATT 2906
Qy 886 ValGlnGlnArgSerProSerThrLysSerAlaSerValThrGlnProLeuGlnHisThr 905
Dh 2907 ATTGAAGAAAGGCTCTCTTCTGATATCTGCTTCAGTACAGAACTCTTGAACAAGTA 2966
Qy 906 AlaGlnGlnAlaMetProProValGlnGlnValThrGlnLysAspLysLysAla 924
Dh 2967 GAAGCTGAAGCCGCACTGTTAACTGAGAGGATTTGAAAGAGAACTTAATGAGAAAGA 3026
Qy 925 GlnThrProValLeuThrGlnThrLeuProGlnGlnLysAspAlaHisAspAspMetVal 944
Dh 3027 GAACCCCAACGCTTACCTGAACTCTGCCAGAGAACAGAGAGGCGCGGGCGCACAGCTC 3086
Qy 945 ThrSerGlnValAspPheThrSerGlnAlaValThrAlaThrGlnThrSerGlnAlaLeu 964
Dh 3087 GTTAGAGAGCGGAATTTGACCCCGAAGCTGTGACAGCTGACAGAACTCAGAGGCCATTG 3146
Qy 965 ArgThrGlnGlnValThrGlnAlaSerGlnAlaGlnGlnThrThrAspMetValSerAla 984
Dh 3147 GGTCGCCAAGAAAGAACCAAGACATCTGCTGGAAGAGACCAACAATAATGGTGTACGA 3206
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGlnGlnAlaThrProValGlnGlnVal 1004
Dh 3207 GTCTCCCAATTAAACGACTCCCAAGACACCAAGAGAGGCCACTCGGTGCGAGAGGTG 3266
Qy 1005 GlnSerGlnValLeuAspThrThrGlnGlnGlnArgGlnThrGlnAlaLeuGlnAla 1024
Dh 3267 GAAGGTGGGTACTCTGACATAGAAAGCAAGAGAGCGGACCTCAAGAGTCTCCAGGCA 3326
Qy 1025 ValAlaAspLysValLysGlnGlnSerGlnValProAlaThr 1038
Dh 3327 GTGGCGAAGAAAGTAAAGAGAAATCCAGCTGCGGACCGGTGGCCAGAAAGTGTG 3386
Qy 1039 ---GlnThrValGlnArgThrGlnSerLysAlaLeuGlnLysValGlnGlnValGlnGln 1057
Dh 3387 CTTCAAGCTGTGCGAGAGCA-----GAGGCAAGAAAGACCAAGAGAG 3428

Qy 1058 AspSerGlnValLeuAlaSerGlnLysGlnLysAspValMetProLysGlnProValGln 1077
Dh 3429 CAGGCTGAAGCGCTCGGCTCTGAAGAAAGACGAGTTGATGTTGAAGATGATGCTCAG 3488
Qy 1078 GlnAlaGlnAlaGlnHisLeuAlaGlnGlnLysSerGlnThrGlnGlnAlaThrProGlnSer 1097
Dh 3489 GAGCGAAAACTGAGCCCTTTACACAGAGAGGTGCGGGGACAGACCCCGCAAGAAC 3548
Qy 1098 LeuGln---ValProGlnValThrAlaAspValAspHis-----ValAlaThrCys 1113
Dh 3549 TTGGAAGAAAGCTCCTCAGTACAGAGACATAGATGCTCAGTGTAACTGACTTGT 3608
Qy 1114 GlnVal-----LysLeuGlnGlnLeu---MetGlnGlnAlaValAla 1127
Dh 3609 CAAGCCGAACCTTAGCTGGGTAAATCACAGGAGATGTGAAATCCATGAGAGAAATGAG 3668
Qy 1128 ProGlnSerSerGlnThrLeuThrAspSerGlnThrAsnGlySerThrProLeuAlaAsp 1147
Dh 3669 CCTGACTCGTGGAAACCCCTTACAGACAGTGAAGTGAATGAGACCCCGTACGCCAGC 3728
Qy 1148 SerAspThrAlaAspGlnThrGlnGlnAspGlnThrLysAspSerGlnAspSerLysAla 1167
Dh 3729 TTGACGACACAGCAGCAACCCAGAAAGACGAGATTTGAAATCCATGAGAGAAATGAG 3788
Qy 1168 ThrAlaAlaValArgGlnSerGlnValThrGlnGlnGlnAlaAlaThrAlaGlnLysGln 1187
Dh 3789 GTGGATCTGTGATCCCACTCAGGCGGCGCAGAAAGCAGAGGCGATTCTGCAAGAAAG 3848
Qy 1188 GlnProSerThrLeuProAsnAsnValProAlaGlnGlnGlnHisGlnGlnLysProGln 1207
Dh 3849 AGGCTTCCAGCA---CTTCCAGTTTGTGTCTCCAGAAAGAACTAAAGAACATCAAG 3905
Qy 1208 ---ArgAspValLeuGlnProThrGlnGlnLeuThrAlaAlaAlaValProValLeu 1226
Dh 3906 ATGAAGACACTCTGACACATACATGAAGAGGTGATGAGTGAAGATCTGATTCCTG 3965
Qy 1227 AlaLysThrGlnValGlnGlnGlnGlnGlnValLysAspThrPheLysAspGlnLysValLys 1246
Dh 3966 TCAAAAGACTGAGGGGACTCA-----GAGGCTGACAGTATGCTGATGAGAAACCAAA 4019
Qy 1247 -----GlnGlnGlnGlnValPheValHisSerGln-----ProAsnSer 1259
Dh 4020 GACGTACCATTTTGAAGAGACTTGAAGGCTTATAGACACAGGATATAACACTGCTGG 4079
Qy 1260 GlnLysAlaAlaAspValThrThrAspSerGlnValMetGlnValAlaGlnCysGlnGln 1279
Dh 4080 GAAAGGTCACTGAAGTGGCTTAAAGTGAAGGAGCAGAAAGAGCTGAATGTAAAG 4139
Qy 1280 LysGlnSerThrGlnValGlnSer-----LeuSerLeuGlnGlnGlnGln 1294
Dh 4140 GATGATGCTCTTGAACCTGACAGTCACTTAAGTCTCTCCATCTCCCGGTGAGAGAG 4199
Qy 1295 MetGlnThrAspValGlnLysGlnLysArgGlnThrLysProGlnGlnValSerGlnGln 1314
Dh 4200 ATGTGATTTCAAGTGAAGAGAGAAACAGAGAGACAGACCAACCATGTGAATGAAG 4259
Qy 1315 Gln---GlnGlnGlnThrAlaAlaProGlnHisGlnGlnLysGlnLysProValLeu 1333
Dh 4260 AAGCTTGAAGCAAGAAACAGCTGTACCGTATCTGAAGAGTCAATGAACAGCTCTCCAG 4319
Qy 1334 ThrLeuAspMetProSerSerGlnArgGlnLysAlaLeuGlnSerLeuGlnGlnGln 1352
Dh 4320 ACAGTGAATGTGCCCATATATAGGCGCAAGAGAGTTCAGAGTTTGAAGAGAGCCCT 4379
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlnCysLysGlnGlnGlnValGlnGlnSerLeuAsp 1372
Dh 4380 CTTCCCTGCTTGGTCAAGAGAGGAGCAAGTATGACCAAAATTTCAAGTTCAAGCTCTGAG 4439
Qy 1373 ThrThrValThrGlnThrAlaGlnAlaValGlnLysValLys-----GlnThrValVal 1390
Dh 4440 GCATCATTCACCTTAACAGCGGCTGCAGAGAGGAAAGGTCTTGAAGAAATCGCAAC 4499

Db 1011 AAGAAGCTGACAGATCTCCGACTAGTCCCGTAGCAGTGAACAGAGATCA 1070
Qy SerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSerPheLys 302
Db 1071 ACCTTCAAAAATTTCTTCACTCAAGGTTGGCGCGCTGGCGCAAAAAGACCAAGTTTCAGG 1130
Qy LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGluGluAlaGluLys 322
Db 1131 AAGCCGAAGAGAGATGAACTGGAAGCTTCAAGAGAAAGAAAGAACAGAACAGCAAGAAAA 1190
Qy 322 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
Db 1191 GTAGACACGAAGAGAACGGAAGAGAGCTTCCCTCCGAGAAACTGACCCGCTCCGAG 1250
Qy 338 -----GluGluProAlaGluAspThrAspGluAlaArgLysSerAlaAspTyr 353
Db 1251 CAAGCCACCCACAGAGAGCCGGCAAGAAAGTGCACAGAGCCCGGTTATCAGCTGAATAT 1310
Qy 354 GluLysValGluLeuProLeuGluAspGluValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1311 GAGAAAGTTGAGCTGCCCTCAGAGAGAGCAAGTCAGTGCCTCCGAGGACCTTCTGAAGAG 1370
Qy 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlu---Glu 392
Db 1371 AAACCTGCTCCGTTGGCGACAGAGCTTTGATGAGAAATAGAACTCCACCAAGAAAGAG 1430
Qy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGluGluGluGly 412
Db 1431 GTGTGGCGGAAGTCCAGCTCAGCACCGTGAGAGAGAAACGAAAGAGCAG----- 1481
Qy 413 GlyGluAlaGluGluGlyValValValGluGluGlyThrGlyGluSerLeuProProGluLys 432
Db 1482 -----AAAAAGAGGTGGAAGAAAGAAAGCAAGGAGCTCTGTGCAGCTGAAGAA 1526
Qy 433 LeuAlaGluProGluGluValProGluGluAlaGluProAlaGluGluLeuMetLysSer 452
Db 1527 TTGGTTGGAATGATGATCAGAACTCAGAAAGCGGAACCTGCCAAGAGAGCTGGTGAAGCTC 1586
Qy 453 ArgGluMetCysValSerGlyGlyAspHisThrGluLeuThrAspLeuSerProGluGlu 472
Db 1587 AAGAAAGCTGTGTTCCTCCGAGAGAGACCTTACACAGAGAGCTCAGCTCATGTAGAG 1646
Qy 473 LysThrLeuProLysHisProGluGluGlyLeuValSerGluValGluMetLysSerGlu 492
Db 1647 AAGGTCTCTCCAAACCCCGGAAGGCGTTGTGAGGTGGAATGCTGTCTATCACAG 1706
Qy 493 GluArgIleLysValGluGluSerProLeuLysLysPheSerSerSerGlyLeuLys 512
Db 1707 GAGAGAAATGAGGTGCAAGGAGAGTCCACTAAAGAGCTTTTACCAAGCTGCTTAAAA 1766
Qy 513 LysLeuSerGlyLysLysGluLysGlyLysArgGlyGlyGlyLysAspGluGluProGly 532
Db 1767 AAGCTTTCTGGAAGAAAGAAAGAAAGGAAAAAG---GAGAGAGAGACGAGAAATCAGGG 1823
Qy 533 GluTyrGluHisIleHisThrGluSerProGluSerAlaAspGluGluLysGlyLysSer 552
Db 1824 GAGCACACTACAGGTTCCAGCGGATCTCCGACAGACAGAGAGAGGAAAAAGGCGAGAGC 1883
Qy 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1884 TCTGCTCTATCCCTTACAGAGCCCGAGAGATCAGCTGTCTGAAAAAGGCTTATGCCGAG 1943
Qy 573 AlaProGluAspGlyLysLysGluGluGlyThrThrSerAspGlyLysLysLysArgGlu 592
Db 1944 GAGCAGACAGAGATGGAGAGCTGAAGAGAGCTACTCCGATGAGAGAAAAAGAGAA 2003
Qy 593 GlyIleThrProThrAlaSerPheLysLysMetValThrProLysLysArgValAlaArg 612
Db 2004 GGTTGACTCTCCCGGCAATCATTCAAAGAGATGTGACCGCCAAAGAGCGTGTAAACGG 2063
Qy 613 ProSerGluSerAspLysGluGluGluGluLysValLysSerAlaThrLeuSerSer 632
Db 2064 CTTTCGAAAGTGAATAAAGATGAGCTGGAAGAGGTCAAGAGCCCTTCTGTCTTCC 2123

Qy 633 ThrAspSerThrValSerGluMetGluAspGluValLysThrValGlyGluGluLys 652
Db 2124 ACCGAGACACAGAGCTCTGGAATGCAAGAGAAAGAAAGGAGCGGTGAAGAGCAAG 2183
Qy 653 ProGluGluProLysArgArgValAspThrSerValSerTTPGluAlaLeuIleCysVal 672
Db 2184 CCGAGAAACCAAGAGCGCAAGGTGATACCTCAGTATCTTGGGAAGCTTAAATTTGTGTG 2243
Qy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyLysProArg 692
Db 2244 GGATCATCAAGAAAGAGCAAGAGAGAGTCTCTTCTGATGAGAGAGGGGAGCAAA 2303
Qy 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2304 GCATGGAGAGAGACACACAGAACTGATAGCGCCGGAAGAAAGCAAGAGCGGGAGCA 2363
Qy 713 AspAlaValProAlaSerThrGluGluGluAspGluAlaGluGlySerSerSerProGlu 732
Db 2364 GAGGGATCTCTTGTCTGTTCACAGAACATGATCCAGGGCAGGGAGAGTTCTCCCGAG 2423
Qy 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTTPGluSerPheLysArgLeu 752
Db 2424 CAAGCTGGAGAGCTTACCGAAGGGAGGGCGTTTCCACTGGAGTCAATTAAGAGTTA 2483
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGluGluValAlaGluAsp----- 769
Db 2484 GTACGCCAAGAAAGAAATCAAAAGTCCAAAGCTGGAAGAGAAAGCAAGACTCCATAGCT 2543
Qy 770 ---SerSerValGluGluLeuSerThrGluLeuGluProSerArgGluGluSerThrVal 788
Db 2544 GGCTGTGTGTAACATTCACCTCAGACACTGAACCCGGTAAAGAAATCTCGGCTC 2603
Qy 788 SerIleLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGluGluGlu 808
Db 2604 TCATCAAGAAAGTTATTCCTGACGAGAAAGAAAGGAGCCAGATGGGAACCAAGACA 2663
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2664 GCCCTGTGTAAGAGCAGGGCCCAAGGGGCCCAAGAAATGATCTCATGTCTCCGGCTC 2723
Qy 828 ValValProLeuSerGlyTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2724 GTGTTCCTCTGTCTGAGTATGATCTGTGAAGAGGAGAAATGAGAGCACAGAGCTC 2783
Qy 846 GlnGlyAsnThrGluLeuProGluLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2784 CAAGAAAGGCGCAGACACCGCAGCAGAAAGCAGCAGCTGAGGTCCAGAGAGCTCAGC 2843
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2844 GAGATCAGGTTCAATATGATGACAGCAGCTGCTGTCAGGAGCAGAGGAGCTACCATT 2903
Qy 886 ValGluGluArgSerProSerThrIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2904 ATTGAAGAAAGGTCTCTCTTGTGATATCTGCTTGAAGACAACTCTTGAACAAGTA 2963
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
Db 2964 GAAGTCGAAGCGCAGCTGTAACTGAGAGGTAATTGAAGAAAGAGATATTCAAGAGA 3023
Qy 925 GluThrProValLeuThrGluThrLeuProGluGlyLysAspAlaHisAspMetVal 944
Db 3024 GAACCCCGCAGCTTACGAACTTCGACAGAGAAACAGAGAGCCGGGGCCACACCGGTC 3083
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3084 GTTATGAGGCGGAATTAACCCCGAAGCTGTGACAGCTGCAAGAACTGCAGAGCCATTG 3143
Qy 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3144 GGTTCGAAGAAAGAAACGAGAGCATCTGCTGTAAGAGACCAAGAAATGCTGTACAGA 3203

QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3204 GTCTCCAGTTAACCGACTCCCCAGACACACAGAGAGGCCACTCGGTGCAGAGGGTG 3263
QY 1005 GluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3264 GAAGGTGGCGTACCTGACATAGAGAGACAGAGAGGGGAGCTCAAGAGGTCTCCAGGCA 3323
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3324 GTGGCAGAAAAAGTGAAGAGGAATCCAGCTGCTGGCACCGCTGGCCAGAGATGTG 3383
QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3384 CTTACGCTGTGCAGAGACA-----GAGGCAGAAAGACCCAGAGAG 3425
QY 1058 AspSerGluValLeuAlaSerGluLysAspValMetProLysGlyProValGln 1077
Db 3426 CAGGCTGAAGCGTGGGTCTGAAGAGAGACCGGATGTAGTGTGAAGTAGATGCTCAG 3485
QY 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
Db 3486 GAGGCAAAAACTGAGCGCTTTACACAGAGGAAGGTGTGGGGCAGACCAACCCAGAAAGC 3545
QY 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3546 TTTGAAAAAGCTCCTCAAGTCACAGAGACATAGAGTCCAGTGAGCTGTGAACCACTTGT 3605
QY 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db 3606 CAAGCCGAACCTTAGCTGGGTAAATATCACAGAGATGTGTGAACAGGTATCTCCC 3665
QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3666 CTTGACTCGGTGGAACCCCTCAGACAGTGAGACTGATGGAAGCACCCTCCGTAGCCGAC 3725
QY 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
Db 3726 TTTGAGCACCAGCGCAACCCAGAAAGACGAGATTGTGGAATCCATGAGAGATGAG 3785
QY 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluAlaAlaThrAlaGlnLysGlu 1187
Db 3786 GTCGATCTGGTACCCAGTCAGGGGACAGAGAGGAGGAGCTTCTGCACAGAAAGAG 3845
QY 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluHisGlyGluGluProGly 1207
Db 3846 AGGCTCCAGCA---CCTCCAGTTTTGTGTTCCAGGAACAACTAAAGAACATCAAG 3902
QY 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu 1226
Db 3903 ATGGAGACACTCTAGAGCATACAGATAAAGAGGTGTCAAGTGGAAACTGTATCCATTCTG 3962
QY 1227 AlaLysThrGluValGlnGluGlyGluValAspThrLeuAspGlyGluLysValLys 1246
Db 3963 TCAAGACTGAGGGAGTCAA-----GAGCTGACCAAGTATCTGATGAGAAACCAAA 4016
QY 1247 -----GluGluGlnGlnValPheValHisSerGly-----ProAsnSer 1259
Db 4017 GAGGTACCATTTTTCGAGGAGCTTGAAGGGTCTATAGACACAGGCATCAACAGTCAGTCGG 4076
QY 1260 GlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4077 GAAAGGTCACTCAAGTTGGCTTAAAGGTGAAGGGACAGAGAGCTGAATGTAAAGAG 4136
QY 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
Db 4137 GATGATGCTCTTGAAGTGCAGAGTCAAGTAAAGTCTCTCCATCCCTCCCGGAGAGAGAG 4196
QY 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
Db 4197 ATGGTAGTTCAAGTCGAAAGGGAGAAACAGAGACAGACCAACCCATGTGAATGAGAG 4256
QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeu 1333

Db 4257 AAGCTTGACACCAACAGCTGTTACCGTATCTGAAGAGAGGTGAGTCAAGCTCTCCAG 4316
QY 1334 ThrLeuAspMetProSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
Db 4317 ACAGTGAATGTGCCATCATAGATGGGCAAGGAAGTCAAGAGTTTGGAGGAAGCCCT 4376
QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4377 CTTCCCTGCTAGGTCAAGAGGAGGAGTATGCACCAAAATTCAGATTTCAGAGCTCTGAG 4436
QY 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
Db 4437 GCATCAATTCATCTAAACAGCGGTGCAGAGGAGGAAGAGTCTTAGGAGAAACTGCCAAC 4496
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4497 ATTTTAGAACAAGGTGAACAGTTGGAGCTGCAGGCTGCACATTTAGTCTTCTGGAAGAGAA 4556
QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4557 TCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGAGAGATGCTGTGCCACAGGG 4616
QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4617 CCGACTGTGAGCAAAATCGACACAGGTAGTAGTATCTGTACTACCAAGAAAGGCTTA 4676
QY 1451 HisProAspLeuGlnGlyLysSerAlaSerGlnArgSerGluGluGluAsp 1470
Db 4677 AGTTCGAGCTGGAGAGAGAGAAACCATCATCTGAAGTGGAGTCAGATGAGTTCGAT 4736
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4737 GAGCAGGTGCTTGCAGCAGGAGTCT-----AAAGTCAGTGTAGCAATTGAGGATTTA--- 4787
QY 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4788 -----GAGCCTCGAAATGCGGATTTTGGAACTTGAGACCAAAAGCAGTAAACTTGTCCA 4841
QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4842 AACATCATTCAGACAGCCGTTGACCAAGTTTGTACGTACAGAGAAACACGCCCGGAATG 4901
QY 1528 HisAlaTyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4902 TTGACGTCTGAGTTACAGACACAGCTCAGCTGATAAAAGCTGACAGCCAGCGCTGGA 4961
QY 1548 ArgCysTrpThrLys-----MetLysAspAla---LysMetLysHisProVa 1562
Db 4962 CAGGAAACGGAGAAAGAGAGAGAGGAACTCAGGCTCTGACAGGATGAAACCAATTT 5021
QY 1562 lProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5022 ACTTCAGCAAGAGAGAGTACAGTCAACCGGAGTGGGAGCAAGACCATTTCTGATATTTCC 5081
QY 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5082 AAAGACATGAGTGAAGCCTCAGAAAAAGACCATGACTGTTGAGGTAGAGGTTCCACTG 5139
RESULT 5
US-09-724-676-12736
; Sequence 12736, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12736
; LENGTH: 5773
; TYPE: DNA

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (91)..(91)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12736

```

Alignment Scores:

```

Pred. No.: 2,33e-162 Length: 5773
Score: 4104.50 Matches: 912
Percent Similarity: 69.40% Conservative: 179
Best Local Similarity: 58.02% Mismatches: 389
Query Match: 50.84% Indels: 92
DB: Gaps: 29

```

US-09-902-432-4 (1-1596) x US-09-724-676-12736 (1-5773)

```

QY 89 ValGlglnArgGluSerGluAspValArgGluLysAspArgValGluGluMetAla1a 108
DB 352 GTTGGACAGAGAGACTCTGAAGATGTGAGCGAAGAGACTCCGATTAAGAGATGGCTACT 411
QY 109 AsnSerThrAlaValGluAspIleThrLysAspGlyGlnGluGluThrSerGluIle1e 128
DB 412 AACTCAGCGGTTTCTTCCAGCATCAGATGATGGCAGAGAGAGACCCGAAATATATC 471
QY 129 GluGlnIleProAlaSerGluAsnAsnValGluGluMetValGlnProAlaGluSerGln 148
DB 472 GAACAGATTCTCTTCCAGAAACCAATTAGAGAGCTAACACACCCTGAGTCCAG 531
QY 149 AlaAsnAspValGlyPheLysLysValPheLysPheValGlyPheLysPheThrValLys 168
DB 532 GCTAAAGATATGGATTAAAGAGGTGTTAAGTTGTTGCTTTAAATTCTACTGTGAA 591
QY 169 LysAspLysAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGly 188
DB 592 AAGGATTAAGACAGAGAGAGCCTGACACTGTCTCACTCTGAGAAAGATGAGAG 651
QY 189 GluGlyAlaGluLysAspValGlyAlaGlyAspHisGlnGluProSerValGluThrAla 208
DB 652 GAGGGAGCA-----GCAGGGGCTGGCGACACCCAGAGACCCCAAGCTT-----GGG 696
QY 209 ValGlglnSerLysSerLysGluSerGluLysGlnSerThrGluLysGlnGluGly 228
DB 697 GCTGGAGAGACAGATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 756
QY 229 ThrLeuLysGlnGluLysSerThrGluIleProLeuGlnAlaGluSerAspGlnAla 248
DB 757 ACCCTGAAGCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 816
QY 249 AlaGluGluGluAlaLysAspGluGlyGluGluLysGlnGluLysGluProThrLysSer 268
DB 817 GTC---GAGGAAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 873
QY 269 ProGluSerProSerSerProValAsnSerGluThrThrSerSerPheLysLysPhePhe 288
DB 874 GCAGATCTCCGACTGATCCCGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
QY 289 ThrHisGlyThrAlaGlyTrpArgLysLysThrSerPheLysSerLysGluAspAsp 308
DB 934 ACTCAAGGTTGGGCGCGCTGGCCCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 993
QY 309 LeuGluThrAlaGluLysArgLysGlnGlnGluAlaGluLysValAspGluGluLys 328
DB 994 GTGAGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1053
QY 329 GluLysThrGluProAlaSerGluGlu-----GlnGlu 339
DB 1054 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1113
QY 340 ProAlaGluAspThrAspGlnAlaGlyLeuSerAlaAspTrpGluLysValGluLeuPro 359
DB 1114 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1173

```

```

QY 360 LeuGluAspGlnValGlyAspLeuGluLysSerGluGluLysCysAlaProLeuAla 379
DB 1174 TCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1233
QY 380 ThrGluValPheAspGluLysMetGluAlaHisGln---GluValAlaAlaGluValHis 398
DB 1234 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1293
QY 399 ValSerThrValGluLysThrGluGluGluGluGluGluGluGluGluGluGluGluGly 418
DB 1294 GTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1329
QY 419 ValValAlaGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGlu 438
DB 1330 ACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
QY 439 ValProGlnGluAlaGluProAlaGluLysLeuMetLysSerArgGluMetCysValSer 458
DB 1390 GAACCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1449
QY 459 GlyLysAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHis 478
DB 1450 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1509
QY 479 ProGluGlyIleValSerGluValGluMetLeuSerSerGlnGluAlaGlyLysValGln 498
DB 1510 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1569
QY 499 GlySerProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLys 518
DB 1570 GGAAGTCCACTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1629
QY 519 GlnLysGlyLysArgGlyGlyGlyLysAspGluGluProGlyGluTrpGlnHisIleHis 538
DB 1630 CAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
QY 539 ThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGlu 558
DB 1687 GCCGATCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1746
QY 559 GluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGlu 578
DB 1747 GAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1806
QY 579 AlaGluGluGlyThrThrSerAspGlyLysLysLysArgGluGlyIleThrProThrAla 598
DB 1807 GCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1866
QY 599 SerPheLysLysMetValThrProLysLysArgValArgProSerGluSerAspLys 618
DB 1867 TCATTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1926
QY 619 GluGluGluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSer 638
DB 1927 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1986
QY 639 GluMetGlnAspGluValLysThrValGlyGluGluGluLysProGluGluLysProLysArg 658
DB 1987 GAAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2046
QY 659 ArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArg 678
DB 2047 AAGGTGAGATCCTCAGTATCTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2106
QY 679 AlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuIleGlyLysAspSer 698
DB 2107 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2166
QY 699 HisArgAlaGluGluLysSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer 718
DB 2167 CAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2226
QY 719 ThrGlnGluGlnAspGlnAlaGlnLysSerSerSerProGluProAlaGlySerProSer 738

```

Db 2227 TCCAGAACATGATCAGGGCAGGGAAGTTCTCCCGGAGCAAGCTGGAAGCCCTTACC 2286
Qy 739 GluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLys 758
Db 2287 GAAGGGAGGGCGTTTCCACCTGGAGTCATTTAAAGGTTAGTCAGCCCAAGAAAAA 2346
Qy 759 SerLysSerLysLeuGluLysAlaGluAsp-----SerSerValGluGln 774
Db 2347 TCAAAGTCCAAGCTGAAGAGAAAGCAAGACTCCATAGCTGGGTCTGGTAGAACAT 2406
Qy 775 LeuSerThrGluLeuGluProSerArgGluGluSerTrpValSerIleLysLysPheIle 794
Db 2407 TCCACTCCAGACACTGAACCCGTAAGAAGAAATCTGGGGTCTCAATCAAGAAGTTATT 2466
Qy 795 ProGlyArgArgLysLysArgAlaAspGlyLysGlnGluAlaThrValGluAspSer 814
Db 2467 CTGGAGCAGGAAGAAAGCCAGATGGGAACACAGCAAGCCCTGTTGAAGACGCA 2526
Qy 815 GlyProValGluIleLeuGluAspProAsnValProAlaValValProLeuSerGlu 834
Db 2527 GGCCCAACAGGGCCCAACGGAAGATGACTCTGATGTCCCGCCGCTGCTCTGTGAG 2586
Qy 835 TyrAsnAlaValGluArgGluLysMetGlu-----AlaGlnGlyAsnThrGluLeu 851
Db 2587 TAGTATGCTGTAGAAAGGGAGAAATGGAGGCACAGCAAGCCCAAAAGCCGAGAGCAG 2646
Qy 852 ProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThr 871
Db 2647 CCCGACGAGAAGCCAGCCACTGAGGTGTCCAAGAGCTCAGCAGAGTCAGGTTTCATATG 2706
Qy 872 ValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSerPro 891
Db 2707 ATGGCAGCAGCTTCGCTGACGGGACGAGGCGAGCTACCATATTATGAAGAAGGTCCTCT 2766
Qy 892 SerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro 911
Db 2767 TCTTGATATCTGCTCAGTGCAGAACCTCTTGAACAAGTGAAGCTGAAGCCGCACGTG 2826
Qy 912 ProValGluGluValThrGluLysAspIleAla-----GluGluThrProValLeuThr 930
Db 2827 TTAACGTGAGGAGTATGGAAAGAGAGTAAATTCGAGAAGAAACCCCCACCGTTACT 2886
Qy 931 GlnThrLeuProGluGlyLysAspAlaHisAspMetValThrSerGluValAspPhe 950
Db 2887 GAACCTCTGCCAGACAACAGAGAGGCCCGGGCGCACACGCTCGTTAGTGAGCGGAATTG 2946
Qy 951 ThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThr 970
Db 2947 ACCCCCAAGCTGTGACAGCTGCAGAAACTGCAGGGCCATTGGGTGCCGAAGGAACCC 3006
Qy 971 GluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAsp 990
Db 3007 GAAGCATCTGCTGCTGAGAGACACACAGAAATGGTGTACAGATCTCCCAAGTTAACCGAC 3066
Qy 991 SerProAspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAsp 1010
Db 3067 TCCCCAGACACCACAGAGGAGGCCACTCCGGTGCAGGAGTGAAGGTGCGTACCTGAC 3126
Qy 1011 ThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLys 1030
Db 3127 ATAGAAGACAGAGAGCGCGACTCAAGAGGTCTCCAGGAGTCCAGGAGTGGCAGAAAAAGTGAA 3186
Qy 1031 GluGluSerGlnValProAlaThr-----GlnThrValGlnArg 1043
Db 3187 GAGAAATCCAGCTGCTCGCCACCGGTGGGCCAGAGATGTGCTTACGCTGTGCAGAGA 3246
Qy 1044 ThrGlySerLysAlaLeuGluLysValGluGluValGluAspSerGluValLeuAla 1063
Db 3247 GCA-----GAGGCAGAAAGACCAAGAGCAGGCTGAGCGTCGGGT 3288
Qy 1064 SerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHis 1083
:::||||| |||||::: |||||

Db 3289 CTGAAGAAAAGAGAGCGATGTAGTTGAAAGTAGATGCTCAGGAGCAAAAACTGAGCCT 3348
Qy 1084 LeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGlu---ValProGlu 1102
Db 3349 TTTACACAAGGGAAGGTGTGGGGCAGACCACCCCAAGAAAGCTTTGAAAAAGCTCTCAA 3408
Qy 1103 ValThrAlaAspValAspHis-----ValAlaThrCysGlnVal----- 1115
Db 3409 GTCACAGAGCATAGAGTCCAGTGAGCTTGTAAACCACCTTGTCAAGCCGAAACCTTAGCT 3468
Qy 1116 ---IleLysLeuGlnGlnLeu---MetGluGlnAlaValAlaProGluSerSerGluThr 1133
Db 3469 GGGTAAATTCACAGAGATGTTGATGGAAACAGGCTATCCCCCTGACTCGGTGGAAACC 3528
Qy 1134 LeuThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGly 1153
Db 3529 CCTACAGACAGTGAGACTGATGAAGCACCCCGTAGCCGACTTTTCACGCACCAGGCACA 3588
Qy 1154 ThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaValArgGln 1173
Db 3589 ACCAGAAAGACGAGATTGTGAAATCCATGAGGAGAAATGAGTCCGATCTGTTACCCAG 3648
Qy 1174 SerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluProSerThrLeuPro 1193
Db 3649 TCAGGGGCGACAGACAGAGGAGTTCCTGCACAGAAAGAGGCTCCAGCA---CCT 3705
Qy 1194 AsnAsnValProAlaGlnGluHisGlyGluGluProGly---ArgAspValLeuGlu 1212
Db 3706 TCCAGTTTGTGTTCCAGGAAGAACTAAGAACAACTCAAGATGGAAGACACTCTAGAG 3765
Qy 1213 ProThrGlnGlnGluLeuThrAlaAlaValProValLeuAlaLysThrGluValGly 1232
Db 3766 CATACAGATAAAGAGGTGTCAGTGGAAACTGTATCCATTTCTGTCAAAGACTGAGGGGACT 3825
Qy 1233 GlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys-----Glu 1247
Db 3826 CAA-----GAGCTCACCATGATGCTGATGAGAAACCAAGACACTCATTTCGAA 3879
Qy 1248 GluGlnGluValPheValHisSerGly-----ProAsnSerGlnLysAlaAlaAspVal 1265
Db 3880 GGNCTTGGGGGTCTATACACAGACGATCAACGTCAGTCGGGAAAGGTCTACTGAAGTT 3939
Qy 1266 ThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGlyLysGluSerThrGluVal 1285
Db 3940 GCCTTTAAAGGTGAAGGACAGAAAGCTGAATGTATAAAGAGTATGCTCTTGAAGTCG 3999
Qy 1286 GlnSer-----LeuSerLeuGluGluGlyGluMetGluThrAspValGlu 1300
Db 4000 CAGAGTCACGCTAAGTCTCCTCATCCCGTGGAGAGAGATGTTCAAGTCGAA 4059
Qy 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGlyGly---GluGlnGluThr 1319
Db 4060 AGGAGAAACAGACAGACAGCCCATGTAATGAATGAAGAGAGCTTGAGCAGAAACA 4119
Qy 1320 AlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSer 1339
Db 4120 GCTGTACCGTATCTCAAGAGGTCAAGTGAAGCAGCTCTCCAGACAGTGAATGTGCCATC 4179
Qy 1340 SerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer---ProSerLeuProAspGln 1358
Db 4180 ATAGATGGGGCAAGGAAGTCAGCAGTTTGGAAAGGAAGCCCTCTCCCTGCTTAGTTCAA 4239
Qy 1359 AspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThr 1378
Db 4240 GAGGAGCAGTATGCACCAAAATTCAGTTCAAGCTCTGAGGCATCTTCACTCTAACA 4299
Qy 1379 AlaGluAlaValGluLysValIle-----GluThrValValIleSerGluThrGlyGlu 1396
Db 4300 CGCGCTGCAGAGGAGGAAAGGTCTTAGGAGAAACTGCCAACATTTTAGAAACAGGTGAA 4359
Qy 1397 SerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGly 1416
Db 4360 ACCTTGAGCCTGCAGGTGCACATTTAGTTCTGGAAGAGAAATCCTCTGAAAAAATGAA 4419
:::||||| |||||::: |||||

Qy	1417	HiStrpPhrLeuGlnHiSaIaGluAspThVaIProLeuGluYProCiuSerGlnIaGlu	1435
Db	4420	GACCTTGGCCGCTCATCCAGGGGAAGATGCTGTGCCCAAGGGCCCACTGTGAGGCAAAA	4475
Qy	1437	SerIlePcoIleIleValIthnProAlaProGluSerThrLeuHisProAspLeuGlnGly	1456
Db	4480	TGACACCCACAGGATAGTATCTTGCCTACTACCAAGAAAGCTTAAGTTCGACCTCGAAGGA	4533
Qy	1457	GluIleSerAlaSerGlnIaYArgIuaYSerSerGlnGluGluAspIysProAspAlaGlyPro	1476
Db	4540	GAGAAACCAATCATCATGAAATGGAAAGTCAAGATGAAGTCATGAGCAGGCTTGTCCAG	4595
Qy	1477	AspAlaAspGlyIysGluSerThrAlaIleGluIaValLeuIaValIaGluProGluIle	1496
Db	4600	GAGGTC-----AAAGTAGCTGATGCAATTTAGAGGATTTAAGAGCTGAAAATGGGATT	4653
Qy	1497	LeuGluLeuGluSerIysSerAsnIysIleValLeuAsnValIleGlnThrAlaValAsp	1516
Db	4654	TTGGAACTTGAGACCAAAAGCAGTAACTTGTCGCAAAAACATCATCCAGACAGCCGTGAC	4711
Qy	1517	GlnPheAlaArgThr--GluThrAlaProGluIuThrHisAlaIaYAspSerGlnIthGln	1533
Db	4714	CAGTTTGATCTGATCAAGAAAGAACCCACCAAAATGTGACGTCTGATTAACAGACACA	4772
Qy	1536	ValProAlaCySAcGluAspSerAspGluProAsnArgCysThrIlys-----	1552
Db	4774	GCTCAGCTGATMAAAAGCTGACAGCCACAGACGCTGACAGCAAGAACGAGAAAGAGAGG	4833
Qy	1553	-----MethylsAspAla-LysMethylsIAspValProGlnProArgGlu-----	1567
Db	4834	GAACTCTAGGGCTCTGCAACAGATGAAACCAATTACTTCAAGCAAGAGAGAGTCAAG	4893
Qy	1568	-----AspleuGlnValIleuThrValIleuGlnAlaIlePrlaGlnProArgly	1583
Db	4894	TCAACCGCAGTGGAGACAGACATTCGATATTTCMAAAGATGAGTGAAGCTTCAGAA	4953
Qy	1583	bcYsleuProaRgLeuGlnIaValIaProlaProval 1594	
Db	4954	AAGACCATGACTGTTGAGGTGAAAGTTCACATG 4987	
RESULT 6			
US-09-724-676A-12736			
/ Sequence 12736, Application US/09724676A			
/ GENERAL INFORMATION:			
/ APPLICANT: CompuGen LTD			
/ TITLE OF INVENTION: Variants of alternative splicing			
/ FILE REFERENCE: 129181.4 CompuGen			
/ CURRENT APPLICATION NUMBER: US/09/724,676A			
/ CURRENT FILING DATE: 2000-11-28			
/ NUMBER OF SEQ ID NOS: 97222			
/ SOFTWARE: PatentIn version 3.2			
/ SEQ ID NO 12736			
/ LENGTH: 5773			
/ TYPE: DNA			
/ ORGANISM: Homo sapiens			
/ FEATURE:			
/ NAME/KEY: misc feature			
/ LOCATION: (91)_(91)			
/ OTHER INFORMATION: n is a,c,g, or t			
US-09-724-676A-12736			
Alignment Scores:			
Pred. No.:		2,33e-162	Length: 5773
Score:		4104.50	Matches: 912
Percent Similarity:		69.40%	Conservative: 179
Best Local Similarity:		58.02%	Mismatches: 389
Query Match:		50.84%	Indels: 92
DB:		5	Gaps: 29
US-09-902-432-4 (1-1596) x US-09-724-676A-12736 (1-5773)			
Qy	89	ValGlyGlnArgGluSerGluAspValArgGluYAspArgValGluGluMetAlaAla	108

Db 1390 GAACCTCAGGAGCGCAACCTGCCAAGAGCTGGTGAAGCTCAAGAAACGTTGTTTCC 1449
Qy 459 GlycylAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHis 478
Db 1450 GGAGGAGACCTACACAGGAGCTGACCTCAGTCTGATGAGAAGGTGCTGTCCAAACCC 1509
Qy 479 ProGluGlyLeuValSerGluValGluMetLeuSerSerGlnGluArgGlyLeuValGln 498
Db 1510 CCCGAAGCGTTGTGATGAGGTGGAAATGCTGTCATCACAGGAGAGAATGAAGGTGCAG 1569
Qy 499 GlySerProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLys 518
Db 1570 GGAAGTCCACTAAAGAAGCTTTTACCAGCACTGGCTTAAAGAGCTTTTGGAAAGAA 1629
Qy 519 GlnLysGlyLysArgGlyGlyGlyAspGluGluProGlyGluLysGlnHisIleHis 538
Db 1630 CAGAAAGGGAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
Qy 539 ThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerProGlu 558
Db 1687 GCCGATTCTCCGACAGCCAGGAGAGCAAAAGGGCGAGAGCTCTGCCTCATCCCTGAG 1746
Qy 559 GluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGlu 578
Db 1747 GAGCCCGAGAGATCACTGTCTGGAAGAGGCTTACCCGAGGTGCAGAGATGGGAA 1806
Qy 579 AlaGluGluGlyThrThrSerAspGlyGlyLysLysArgGluGlyIleThrProTrpAla 598
Db 1807 GCTGAAGAAGGAGTACTTCGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1866
Qy 599 SerPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLys 618
Db 1867 TCATTCAAAAGATGGTGAGCCCAAGAGCGTGTAGAGCGGCTTCGGAAGGTGATAAA 1926
Qy 619 GluGluGluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSer 638
Db 1927 GAAGATGAGTGGCAAGGTCAAGAGCGTACCTTGTCTCCACCGAGAGACAGCCTCT 1986
Qy 639 GluMetGlnAspGluValLysThrValGlyGluGlnLysProGluGluProLysArg 658
Db 1987 GAAATGCAAGAAGAAATGAAAGGAGCGTGGAGAGCCAAAGCGGAGAGAACCAAGCGC 2046
Qy 659 ArgValAspThrSerValSerTrpGluAlaLeuLysCysValGlySerSerLysArg 678
Db 2047 AAGTGTGATACCTCAGTATCTTGGGAAGCTTAAATTTGTGTGGATCATCAAGAAAGA 2106
Qy 679 AlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSer 698
Db 2107 GCAAGGAGAGGTCCTCTCTGATGAGGAAGGGGACCAAAAGCAATGGGAGGAGACCAC 2166
Qy 699 HisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer 718
Db 2167 CAGAAAGCTGATGAGCGCGGAAAGACAAAGAGACGGGACAGACGGGATCCTTGCTGGT 2226
Qy 719 ThrGlnGluGlnAspGlnAlaGlnGlySerSerProGluProAlaGlySerProSer 738
Db 2227 TCCCAAGAACATGATCCAGGACAGGAAGTCTCTCCCGGAGCAAGTGGGAAGCCCTACC 2286
Qy 739 GluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLys 758
Db 2287 GAAGGGAGGCGGCTTCCACTGGGATGATTTAAAGGTAGTACGCCCAAGAAAAAAA 2346
Qy 759 SerLysSerLysLeuGluGluLysAlaGluAsp-----SerSerValGluGln 774
Db 2347 TCAAGTCCAAAGCTGGAAGAGAGAAAGAGAGACTCCATAGCTGGGTCTGGTGTAGAACAT 2406
Qy 775 LeuSerThrGluLeuGluProSerArgGluGluSerTrpValSerIleLysLysPheIle 794
Db 2407 TCCACTCCAGACACTGAACCGGTAAAGAGATCTCTGGGTCTCAATCAAGAGATTTATT 2466
Qy 795 ProGlyArgArgLysLysArgAlaAspGlyGlnGluGlnAlaThrValGluAspSer 814
Db 2467 CCTGGACGAAGGAAGAAAGGCCAGATGGGAAACAAGAAACAAAGCCCTGTTTGAAGACGCA 2526

Qy 815 GlyProValGluIleLeuLeuAspAspProAsnValProAlaValValProLeuSerGlu 834
Db 2527 GGGCCAAACAGGGGCCAACGAAGATGACTCTGATGTCCGCGCGTGTCTCTGTCTGAG 2586
Qy 835 TyrAsnAlaValGluArgGluLysMetGlu-----AlaGlnGlyAsnThrGluLeu 851
Db 2587 TATGATGCTGTAGAAAGGGAGAAAAATGGAGGCACAGCAAGCCCAAAAGCGCAGAGCAG 2646
Qy 852 ProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThr 871
Db 2647 CCGAGCAGAGAGCAGCCACTGAGGTGTCCAAGAGCTCAGCGAGAGTCAAGTTCATATG 2706
Qy 872 ValSerValAlaValLysAspGlyThrArgAlaValThrSerValGluGluArgSerPro 891
Db 2707 ATGCGCAGCAGCTCGCTGACGGCAGCGAGGCAGTACCATTATTGAAGAAGGTCTCCT 2766
Qy 892 SerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro 911
Db 2767 TCTTGATATCTCTTCAAGTGACAGAACCTTTTGAACAAAGTGAAGCTGAAGCGCGCACTG 2826
Qy 912 ProValGluGluValThrGluLysAspIleAla---GluGluThrProValLeuThr 930
Db 2827 TTAAGTGGAGAGTATTTGGAAGAGAGTAATTGCAAGAGAGAACCCCCACCGTTACT 2886
Qy 931 GlnThrLeuProGluGlyLysAspAlaHisAspMetValThrSerGluValAspPhe 950
Db 2887 GAACCTCTGCACAGAAACAGAGAGGCCCGGGCGCACACGCTCTGTAGTGGCGGAATTG 2946
Qy 951 ThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThr 970
Db 2947 ACCCCCGAAGCTGTGACAGCTGCAGAAACTGCGAGGCCATTTGGGTGCGAAGAGGAACC 3006
Qy 971 GluAlaSerGlyAlaGluGluThrAspMetValSerAlaValSerGlnLeuThrAsp 990
Db 3007 GAAGCATCTCTGCTGAGAGAGACCAAGAAATGTTGTCAGCAGTCTCCAGTTAACCGAC 3066
Qy 991 SerProAspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAsp 1010
Db 3067 TCCCCAGACACACAGAGGAGGCCACTCCGGTGCAGAGGTGGAAGTGGCGTACCTGAC 3126
Qy 1011 ThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLys 1030
Db 3127 ATAGAAGACAGAGAGCGGACTCAAGAGTCTCCAGGAGTGGCAGAAAAAGTGA 3186
Qy 1031 GluGluSerGlnValProAlaThr-----GlnThrValGlnArg 1043
Db 3187 GAGGAATCCAGCTGCTGCGCCGCGTGGCCAGAGAGATGTGCTTCAAGCTGTGCAGAGA 3246
Qy 1044 ThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAla 1063
Db 3247 GCA-----GAGCGAGAAAGACAGAGAGCGAGCGTGAAGCGTCGGGT 3288
Qy 1064 SerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHis 1083
Db 3289 CTGAAGAAACAGAGAGCGATGTAGTGTGAAAGTAGATGCTCAGGAGGCAAAAGTGGAGCCT 3348
Qy 1084 LeuAlaGluGlySerGluThrGlyGlnAlaThrProGluSerLeuGlu---ValProGlu 1102
Db 3349 TTTACCAAGGAGGTGGTGGGGCAGACCCCAAGAAAGCTTTGAAAAAGTCTCTCAA 3408
Qy 1103 ValThrAlaAspValAspHis-----ValAlaThrCysGlnVal----- 1115
Db 3409 GTCAAGAGAGATAGAGTCCAGTCCAGTGTGTAACCACTGTCAAGCCGAAACCTTAGCT 3468
Qy 1116 ---IleLysLeuGlnLeu---MetGluGlnAlaValAlaProGluSerSerGluThr 1133
Db 3469 GGGGTAAAAATCAAGAGATGTGTGATGGAACAGGCTATCCCCCTGACTCGGTGGAAACC 3528
Qy 1134 LeuThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGly 1153
Db 3529 CCTACAGACAGTGAAGTGTGATGGAAGACCCCGGTAGCCGACTTTTGACCGCAGGCACA 3588


```

QY 1154 ThrGlnGlnAspGluThrIleAspSerGlnAspSerIleValaThrAlaValaArgin 1173
Db 3589 ACCCGAAGAGAGAGATGTTGAAATCCATGAGAGAAATGAGCTGCTGTCATCCCG 3648
QY 1174 SerGlnValaThrGlnGlnGlnValaThrAlaGlnIleGlnGluProSerThrLeuPro 1193
Db 3649 TCAGGGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3705
QY 1194 AsnAsnValaProAlaGlnGlnGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGln 1212
Db 3706 TCACATTTTGTGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3765
QY 1213 ProThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1232
Db 3766 CATACAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3825
QY 1233 GlnGlnGlnGlnValaAspIleAspIleAspIleAspIleAspIleAspIleAspIle 1247
Db 3826 CAA-----GAGGCTGACAGATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3879
QY 1248 GlnGlnGlnValaPheValaHisSerGly-----ProAsnSerGlnIleAlaAspVal 1265
Db 3880 GGACTTGAAGGGCTTATAGACACAGCATACAGCATACAGCATACAGCATACAGCATACAG 3939
QY 1266 ThrIleAspSerGlnValaMetGlyValaAlaGlyCysGlnGlnGlnGlnGlnGlnGln 1285
Db 3940 GCCCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3999
QY 1286 GlnSer-----LeuSerLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1300
Db 4000 CAGAGTCAAGCTTAAGCTTCTCCATCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4059
QY 1301 LysGlnIleAspGlnThrIleAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1319
Db 4060 AGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4119
QY 1320 AlaAlaProGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1339
Db 4120 GCTGTACCGTATCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4179
QY 1340 SerGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1358
Db 4180 ATAGATGGGGCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4239
QY 1359 AspIleAspIleCysIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1378
Db 4240 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4299
QY 1379 AlaGlnAlaValaGlnIleValaIle-----GlnThrValaIleSerGlnGlnGln 1396
Db 4300 GCGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4359
QY 1397 SerProGlnIleValaGlnIleValaIleValaIleValaIleValaIleValaIle 1416
Db 4360 AGCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4419
QY 1417 HisIleThrLeuGlnHisIleAspIleValaProLeuGlnGlnGlnGlnGlnGlnGln 1436
Db 4420 GACTTGGCGCTCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4479
QY 1437 SerIleProIleIleValaThrProAlaProGlnGlnGlnGlnGlnGlnGlnGlnGln 1456
Db 4480 TCAGACACCGATATGATATCTCTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4539
QY 1457 GlnIleSerIleAspSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1476
Db 4540 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4599
QY 1477 AspAlaAspGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1496
Db 4600 GAGGCTC-----AAAGTGAAGTGAAGATTTGAGAGAGAGATTTGAGAGAGATTTGAGAG 1653
QY 1497 LeuGlnLeuGlnSerIleSerIleValaIleValaIleValaIleValaIleValaIle 1516

```

```

Db 4654 TTGGAACCTTGAGACCAAAAGAGATTAACCTTGTCCAAAACATCATCTCAGAGAGAG 4713
QY 1517 GlnPheAlaArgThr-----GlnThrAlaProGlnIleThrHisAlaIleAspSerGlnThrGln 1535
Db 4714 CAGTTTGTACCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4773
QY 1536 ValProAlaCysArgLeuAspSerIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1552
Db 4774 GCTCAGCTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4833
QY 1553 -----MetIleAspAla-LysMetIleHisIleProValaProGlnProArgGln----- 1567
Db 4834 GAACCTCAGGCGCTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4893
QY 1568 -----AspLeuGlnValaLeuThrValaLeuGlnAlaThrAlaGlnProArgGly 1583
Db 4894 TCACCCGAGGTGGGCAAGACATTTCTGATATTTCCAAAAGCATGAGTGAAGCTCAGAGA 4953
QY 1583 GcyleuProArgLeuGlnIleValaProVal 1594
Db 4954 AAGACCATGACTGTGTGAGTGAAGAGTTCCACTG 4987

RESULT 7
US-10-133-937-37
; Sequence 37, Application US/10133937
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
; FILE REFERENCE: 11613.56US01
; CURRENT APPLICATION NUMBER: US/10/133,937
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 9161
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-937-37

Alignment Scores:
Pred. No.: 1,41e-12 Length: 9161
Score: 492.00 Matches: 386
Percent Similarity: 33.47% Conserves: 269
Best Local Similarity: 19.72% Mismatches: 714
Query Match: 6.09% Indels: 589
Gaps: 75

US-09-902-432-4 (1-1596) x US-10-133-937-37 (1-9161)
QY 45 AlaAspProAlaThrIleValaLeu-----ProGlnIleAspGlnGlnGlnGlnGln 59
Db 1399 GCAAACTCTGGGAGAGAAATATCATCCAGATCCTGTTTCTGGAGAACAGACCATGATCAAC 1458
QY 60 SerValaGlnGlnValaAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 73
Db 1459 ATCTCGAAGAGGTGGAGAAAGCTCAACATCTTACACTTTTCGAGACAGAGAGAGAGAGAG 1518
QY 74 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 93
Db 1519 CAAAGAGATTCACAGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1578
QY 94 SerGlnAspValaArgGlnIleAspArgValaGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 113
Db 1579 AGGGCTGATAGCGGAGAAAGT-----CTGAAGCCAGCGGCAAAACCACTTCTTACG 1629
QY 114 GlnAspIleThrIleAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 133

```

Db 1630 AATCCGTGCGCAAGGAGTCAAAAGAAACCCCTGAGGTC-----ACA 1674
 Qy 134 SerGluAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGly 153
 Db 1675 AAAGTGAATCAGCTGGAA-----AAGCCACCAAAAGTTGAAGCAAAAGAA----- 1719
 Qy 154 PheLysLysValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGlu 173
 Db 1720 -----AAGTAAATGGTGAAGAAAGCAAGCCAGTA 1749
 Qy 174 LysSerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAla 193
 Db 1750 AAAACAGAGACCAAACTTCAGTGACTGAAAGGAG----- 1785
 Qy 194 SerValGlyAlaGlyAspHisGlnProSer---ValGluThrAlaValGlyGluSer 212
 Db 1786 -----GTTCCCAAGCAAGAGGCCATCTCCAGTGAAGCGGAGTGGCTGAG--- 1833
 Qy 213 AlaSerLysGluSerGluLeuLysGlnSerThrclulysGlnGluGlyThrLeuLysGln 232
 Db 1834 -----AAGCAAGCCACAGATGTCAAAACCCAAAGCTGCCAAGGAG 1872
 Qy 233 GluGlnSerSerThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGlu 252
 Db 1873 AAGCGGTGAAAAGGAAACAAAGGTAAAGCTGAACAGAAAGAGGAGAAAGAAAG 1932
 Qy 253 AlaLysAspGluGlyGluGlnGlnGlnGluProThrLysSerProGluSerPro 272
 Db 1933 CCAAGAAAGAGTGGCTAAAGAGGAGCAAAACACCTATCAAGAGGAGGAGAAACCA 1992
 Qy 273 SerSerProValAsnSerGluThrThrSerSerPheLysLysPheThrHisGlyTrp 292
 Db 1993 AAA-----AAGGAAGAGGTGAAAAAGAAAGTCAAAAAGAGATC----- 2031
 Qy 293 AlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspAspLeuGluThrAla 312
 Db 2032 -----AAGAAAGAGAGAGAAAGAAAGAACCCCAAGAGAGGTAAAGAAAGAACCCG 2082
 Qy 313 GluLysArgLysGluGlnGluAlaGluLysValAspGluGluGlnLysLysThrGlu 332
 Db 2083 CCAAGGAAGTCAAGAGGAAGTTAAGAG-----GAAGAGAAAGAGGAA 2127
 Qy 333 ProLysSerGluGlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAsp 352
 Db 2128 GTGAAAAGAGAGAAAGCAACCCAAAAAGAAAT-----AAGAGCTCCCTAAAGAC 2181
 Qy 353 TyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGlu 372
 Db 2182 GCAAGAAATCATCTACTCTCTG-----TCTGAAGCA 2214
 Qy 373 GluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGlu 392
 Db 2215 AAAAAACCACTGCTTTAAACCAAAAGTACCCAAAGAGGAGAGTCTGTCAAGAAAGAT 2274
 Qy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGly 412
 Db 2275 TCTGTGCTCCGGAAGCAAGGAGAGAGGGAATAAAAGTCATTAAAGAGGAAGGC 2334
 Qy 413 GlyGluAlaGluGlyGlyValValValGluGlyThrclulysLeuLeuProGluLys 432
 Db 2335 AAGGCCGAGAGGCTGCTGCTGAGCTGTGCGCACTGAGCCACCAAGAGCTGTCATG 2394
 Qy 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSer 452
 Db 2395 GCGGCACTGGAATAGCAGC-----ATTGGCCCTGCCAAAGAACTCGAAGCTGAG 2445
 Qy 453 ArgGluMetCysValSerGlyGlyAspHisThrGln---LeuThrAspLeuSerProGlu 471
 Db 2446 AGGTCCCTTATGTCATCTCTCTGAGGATCTAACCAAGGACTTTGAAGAGTTAAAGGCTGAA 2505
 Qy 472 GluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSer 491
 Db 2506 GAGTGCATGTAACAAAG-----GACATCAAGCCTCAGCTGGAGCTTAATCGAAGAC 2556

Qy 492 GlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeu 511
 Db 2557 GAAGAGAAACTGAAG---GAAACTGAGCCAGTCCAAAGCTACCTCATCCAGAAGGAGAGA 2613
 Qy 512 LysLysLeuSerGlyLysLysGlnLysGlyLysArgGly----- 524
 Db 2614 GAAGTCAACCAAGGTCCTGCCGAGTCCCCTGATGAGGGAATCACTACCACTGAAGGGAG 2673
 Qy 525 GlyGlyLysAspGluProGlyGluTyrGlnHisIleHisThrGluSerProGluSer 544
 Db 2674 GGCAGTGTGAACAGACACCTGAGGAGTGGAGCCCGTCGAGAGCAGGAGGTAGAC--- 2730
 Qy 545 AlaAspGlnLysGlyGluSerSerAlaSerProGluProGluGluThrThr 564
 Db 2731 ---GACATTGAAAAATTTGAAGATGAAGGAGCGCGTTTGAAGAATCTTCAGAGACT--- 2784
 Qy 565 CysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGluGlyThrThr 584
 Db 2785 -----GGAGACTATGAAGAGAGGAGGCAACTGAGGAGGCTGAGGAGCCAGAA 2832
 Qy 585 SerAspGlyLysLysArgGluGlyIleThrProTrpAlaSerPheLysLysMetVal 604
 Db 2833 GAGGATGGGAGGAACAC-----GTA 2853
 Qy 605 ThrProLysLysArgValArgArgProSerGluSerAspLysGluGluLeuGlu--- 623
 Db 2854 TGTGTGAGCGCTCCCAAGCACACCCCACTGAGGATGAGGAAAGTCCCAAGCGGAGGCT 2913
 Qy 624 -----LysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGlu 639
 Db 2914 GATGCATACATCAGGAGAGAGGAGTCTGTGCCCACTGGGATGACCGAGCGGAAGAA 2973
 Qy 640 MetGlnAspGlnValLysThrValGlyGluGlnLysProGluGluProLysArgArg 659
 Db 2974 GACATGATGAGCCATTGAGAAAGAGAGGCTGTAACAACTCTGAAGAGGAGGCTGATGAG 3033
 Qy 660 ValAspThrSer-----ValSerTrpGluAlaLeu 669
 Db 3034 GAGGACAAAGCTCAAGATGCCAGAGAGGAGGAATATGAGCCGGAATAATGGAAGCTGAA 3093
 Qy 670 IleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGly 689
 Db 3094 GACTATGTGATGTGTGTGTCGCAAGGCTGCAAGGCTGGTGTGTCGAGGAGCAGTAT 3153
 Qy 690 Gly-----ProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAla 704
 Db 3154 GGATTCTCCACACACCAACCAAGCAACTAGGAGCCCACTCTCTGGCCGAGAACCTGCA 3213
 Qy 705 SerLysAspLysAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGln 724
 Db 3214 TCTTCAATTCAATGAT-----GAGACTTTACCTGGAGGCTCAGAGGAGGCCACC 3264
 Qy 725 AlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGlyGlyValSer 744
 Db 3265 GCTTCTGATGAGGAGAAATCGAGAA-----GACCAGCTCAGGAATTCATCGCACCTCT 3318
 Qy 745 ThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLysLysLeuGlu 764
 Db 3319 GGCTAC-----ACTCAGTCTACTATTGAG 3342
 Qy 765 GluLysAlaGluAspSerSerValGluGlnLeuSerThr----- 777
 Db 3343 ATATCCAGTGAAGCCACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3402
 Qy 778 -----GluIleGluProSerArgGluSerTrpValSerIleLysLys--- 792
 Db 3403 GAGACCAACAATGAAGAGAGCGAGTCCCTTCTCAGGAATTCGTAATAATATCACCAATAT 3462
 Qy 792 ----- 792
 Db 3463 GAATCTTATTGTTTCTCAGGAATACTCTAAACCTGCTGATGTTTACACCGCTCAACGGA 3522

QY 793 PheileProGluValArgArgLysLysArgAlaAspGlyLysGlnGlnAlaThrValGlu 812
Db 3523 TTTTCTGAAGGATCAAAAACAGATGCCACTGATGGCAAGATTACAAATCTTCACGCTCT 3582
QY AspSerGlyProValGluIleAsnGluAspAspProAsnValProAlaValProLeu 832
Db 3583 ACCATATCACCAACC----- 3597
QY 833 SerGluThrAsnAlaValGluArgGlyMetGluAlaGlnGlyAsnThrGluLeuPro 852
Db 3598 -----TCTTCATGAGAGAACAAATTACAGACATGCT----- 3633
QY 853 GlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThrVal 872
Db 3634 -----TTACGTGATGCTTACTGCTGAGTGAAGAAAGCCAGACAC----- 3672
QY 873 SerValAlaValIleAspGlyThrArgAlaValThrSerValGlu-----GluArg 889
Db 3673 ACTTGGACATCAAAAGATAGCATCTCAGCTGTTTCAAGTGAAGAAAGTCAAGCCATCGAAG 3732
QY 890 SerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr---AlaGlyGlu 908
Db 3733 AGCCCGTCTC---CTAGATCATCTCCACATCAGATCACCCTTGAAGAAAGCCCGGTGGA 3789
QY 909 -----AlaMetProProValGlu---GluValThrGluLysAspIleIle 922
Db 3790 CCTAGTGTGAACTCTCTCTGACGCCCAATGAGATTAAAGTCTCTGACAGAGCAGAAAGTA 3849
QY 923 AlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAsp 942
Db 3850 GCGCCGGTGTCTCTGAGGTGACCAAGAGTAAAGTAAAGAAATTTGCTAGTCTGCTGAG 3909
QY 943 MetValThrSerGluValAspPheThrSerGluAlaValThrAlaThr----- 958
Db 3910 GACAAGACTCTGAAAGTGTGTACCATCTCAGTCCGTGACTGGCAGTCTGCTGACACACA 3969
QY 959 -----GluThrSerGluAlaLeuArgThrGluGluVal--- 969
Db 3970 CCTTACTATCAATCTCTTCTGACGAGAAATCCAGTATCTCTTACAGAAAGTCAATTGAA 4029
QY 969 ----- 969
Db 4030 AAACCAACAGCAGTTCACAGTGTGTTGAATTCAGTATGCCAAAGATGAAATGAAG 4089
QY 969 ----- 969
Db 4090 GCTTCAGTAAAGCCCATGATGAGCCGTCGACTCAGAGTCTCTTATGAAAAAGTT 4149
QY 969 ----- 969
Db 4150 TTGTCTCTTATAGCAGCCGCCCTCATTTGATCCGAGTCTGCTTATGAAATTTCTTA 4209
QY 970 -----ThrGluAlaSerGly 974
Db 4210 AGTCTGATGACAGAGCTTCTGGCAGAGGTGCCGAAAGTCTTTTGAAGAAAAAGTGA 4269
QY 975 AlaGluGluThrThrAspMetValSerAlaValSerGluLeuThrAspSer----- 991
Db 4270 AAACCAAGGCTCTCCAGACCAAGTAAGTTCGAATGACTTCTTCAATGACTCTTAC 4329
QY 992 ProAspThrThrGlu-----GluAlaThrProValGlnGluValGluSerGly 1007
Db 4330 CAAGACAAACAGAGAGAGAAAGACAGACACTTGGCAATAAAGAAAGACACTTGGCCAA 4389
QY 1008 ValLeuAspThrGlnGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAsp 1027
Db 4390 GAAAAAGAAACGTGATGATTTGAAGCCATGAGTTCAACACAGCACTGGCTGAGTGA 4449
QY 1028 Lys----- 1028
Db 4450 AGGAATTTAGAGATGTTTCTCCACAAATAGATGTCAGTCAAGTTGATTTTAA 4509
QY 1039 -----ValLysGluGluSerGlnValProAla 1037

Db 4510 GAAGACATAAGATGTCATTTCTGAAAGTACTGTCTCAGCAAGTCCGACTCTCTT 4569
QY 1038 ThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 4570 GATGAGGCGGTAGAGAGAAAGACAGCACTCTCATATGAGGGGTGGCCCTGATCCACA 4629
QY 1058 AspSerGluValLeuAlaSerGluLysGlu-----LysAspValMetProLysGly 1074
Db 4630 GCCTCAGGGCTACAGAGCTCAATTCAGAGCCAAACACAGATGATGTCTCCATCTCTG 4689
QY 1075 ProValGlnGluAlaGlyAlaGluHis----- 1083
Db 4690 CATGCT---GAGGTGGCTCTCCACATTCACAGAAAGTATGACTCCCTTCAGTGTCT 4746
QY 1084 LeuAlaGlnGlySerGluThrGlyGlnAlaThr----- 1094
Db 4747 GTTGTGCAAAACCTACACATTCACAGAAACAGAAATGTCTCCATTAAGAAAGATGC 4806
QY 1095 -----ProGluSerLeuGluValProGluVal-----Thr 1104
Db 4807 CCAGACCGATGTCATATTTCTCCACAGATTTCTCCCTAAACTGCAAAAGTCCAGACACA 4866
QY 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeu-----GlnGlnLeu 1121
Db 4867 CCGTTCAAGATCACAGATCTGAACAGTCTCTCAATGCTTATGAATTTGGCCAGAAATCT 4926
QY 1122 MetGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThr----- 1139
Db 4927 CCGTGAACAATCCCTTGTATGAGACTTCAGACGTCCAGACGTCCAGATCACCCATGCTGT 4986
QY 1140 -----AsnGlySerThrProLeu-----AlaAspSerAsp 1149
Db 4987 GCAGGCGTCTTCATCATCTGAACAAATGGCCAACTGAAGAGGACTGACATGCTCTTAC 5046
QY 1150 ThrAlaAspGlyThr-----GlnGlnAspGluThrIleAspSer 1163
Db 5047 ATGACGAGCTCCAGTTTATTCACATTAAGTACCACTATGAGAGCCCTCTACACCCAA 5106
QY 1164 AspSerLysAlaThrAlaAlaValArgLysSerGlnValThrGluGluGluAlaAlaThr 1183
Db 5107 GATATGATCTTTCTGAGCTCATCTCAGTATCTCAAGTA---GAGGCTCCCGCTCCAC 5163
QY 1184 AlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGly 1203
Db 5164 TCTTCTGCTCATVACCCCTTCTCAGATCCCTCT-----CCTCTCCAA----- 5205
QY 1204 GlnGluProGluArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaVal 1223
Db 5206 GAAGATCTTATCCGATGTGCTCTCCAGAGTATGTCCTTATATGCTTCACTCAC 5265
QY 1224 ProValLeuAlaLysThrGluValGlyGlnGluGlyValAspTrpLeuAspGlyGlu 1243
Db 5266 -----TCTGAAAAGTCCAAAGCTGGAAAGAGAG 5295
QY 1244 LysValLysGluGlnGlnValPheValHisSerGlyProAsnSerGlnLysAlaAla 1263
Db 5296 AAGCTCTCTCCAAATCTGAATATCTCTCAGTACCCACAGAGAGTCTCTCTTATAT 5355
QY 1264 AspValThrThrAspSerGluValMetGlyVal-----AlaGlyCysGlnGlu 1279
Db 5356 TCACCTACTTCTTGAATTTCACTCTGACAGAAAGAGAAAGACGAACTTGGCCACAGT 5415
QY 1280 LysGluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspVal 1299
Db 5416 TCCTCTTCTCCACCAATAGATGACGATCCGAGAGCCCTATGCTTCCGCTTCAGTG 5475
QY 1299 ----- 1299
Db 5476 TTATTCATACATGCAACACCATCTAGCTTGAATAGATTGTCCACACTGGCTG 5535
QY 1300 GlnLysGluLysArgGluThrLysProGluGlnValSer-----GluGluGluGlu 1316

Db 5536 GAGAAGGACAGTGAGGGAAGACACACCTGGTGACTTTTAGCTATCGCTATCAAAAGCCTGAG 5595
 Qy 1317 GlnGluThrAlaAlaProGluHisGluGly-----ThrTyrGlyLysProVal 1332
 Db 5596 GAAACACACAGGTCCTCCAGATGAAGAGATTATGACTATGAGTCTTATGATGAGACACCC 5655
 Qy 1333 LeuThrLeuAspMet-----ProSer 1339
 Db 5656 CGACCTCAGATGGTGGTGCTATTACTATGAGAAGATGAGAGAACCAAAATCTCCA 5715
 Qy 1340 SerGluArgGlyLysAlaLeuGlySerLeuGlySerProSerLeuProAspGlnAsp 1359
 Db 5716 AGTGACAGTGGCTACTCTATGAGACCATTTGGGAAGAACTACCAAGACCCCTGAAGTGT 5775
 Qy 1360 LysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAla 1379
 Db 5776 -----GACTATTCCTAT 5787
 Qy 1380 GluAlaValGluLysValIleGluThrVal-----ValIle 1391
 Db 5788 GAAATTATTGAGAAGACCAACCGGACCCCTGAAGAGGTGGGTACTCATATGACATAAGT 5847
 Qy 1392 SerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSer 1411
 Db 5848 GAAAAGACCAACCGCCCCCGAAGTGAAGTGTACAGCTATGAAAGACTGAGAGTCT 5907
 Qy 1412 -----SerAlaThrGlyHisTrp 1418
 Db 5908 AGAAGGCTTCTGGATGACATCAGCAATGGCTATGATGACTCTGAGGTGGTGGCCAC--- 5964
 Qy 1419 ThrLeu-----GlnHisAlaGluAspThrValProLeuGlyProGlu 1432
 Db 5965 ACACCTGGGAGCCCGAGCTACTCTTATGAAACCACTGAGAAATACCAGATTTCCCTGAG 6024
 Qy 1433 SerGlnAlaGluSerIleProIle----- 1441
 Db 6025 TCTGAGGTATTCTTATGAGATCTTACAAAGACACACGACCCCTGATCTTCCACA 6084
 Qy 1442 -----ValThrProAlaProGluSerThrLeuHisProAsp 1453
 Db 6085 TACTGTTACGAGCTGCAGAGAAAATCACTAGAACCCCTCAGGCATCCACATATCTCTAC 6144
 Qy 1454 LeuGlnGlyLysSerAlaSerGlnArgGluArgSerGluGluAspLysProAsp 1473
 Db 6145 GAGACTTCAGACCTATCTACTACTCGAGAAAAGAAAGTCCCTCCAGAAAGCCGCTCAGAT 6204
 Qy 1474 AlaGly-----ProAspAlaAspGlyLysGlu 1482
 Db 6205 GTCGATTATGCTCGTGCTCTTGTGTAATACAGCACCCCAAGACAGAGCTTTCACCC 6264
 Qy 1483 Ser-ThrAlaIleGluLysValLeu----- 1490
 Db 6265 TCITTCATTAAATCCCAATCCTCTTGTAGTGGTTGGCCAGTGAAGAACCCACTGAAGAATCT 6324
 Qy 1491 -----LysAlaGluProGluIleLeuGluLeuGluSerLysSerAs 1504
 Db 6325 GAAAAGCCCTCACTCAATCAGGGGAGCCCGCTCCAGGAGGAAAGCAACAGGCG 6384
 Qy 1504 nLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThr-- 1523
 Db 6385 CGACAGTGTGATAAACCCTCC-ACCTCAGTCAGCGAGTCAGCCCATCCAGAGCCGA 6443
 Qy 1524 -----AlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysAr 1540
 Db 6444 CTCTGATGTTCCCGGAGATGAAGTGCCCTTCATCAG- -GCCGATGCCAA 6497
 Qy 1540 gLeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHi 1560
 Db 6498 TATCGACTCTGAAGACAGTTCGGAACCACTCCCAAGACCAAACTGTCTACGTACAAACA 6557
 Qy 1560 s-----ProVal-----ProGlnProArgGlu----- 1567
 Db 6558 CATGGACCCACCTCCAGCTCCCGTGCAAGACCGCAGCCCTTCGCCACCGCCCTGATGT 6617

Qy 1568 -----AspLeuGlnValLeuThrValLeuGlu-AlaTrpAlaGln----- 1580
 Db 6618 GTCCATGTGGACCCAGAGCCCTTGGCCATTGAGCAGAAACCTGGGCAAGCTCTAAAGAA 6677
 Qy 1581 -----ProArgLysCysLeuProArgLeuGlnLeuLysAlaPro 1593
 Db 6678 AGATCTGAAAGAGAGACCAACCAAAAGCCAGGTACAAAGACCA 6724

RESULT 8

US-10-203-138A-6112
 ; Sequence 6112, Application US/10203138A
 ; GENERAL INFORMATION:
 ; APPLICANT: Molecular Dynamics, Inc.
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wenheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: PB 0004 WO 8
 ; CURRENT APPLICATION NUMBER: US/10/203,138A
 ; CURRENT FILING DATE: 2002-08-02
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 04 February 2000 (04.02.00)
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 26 May 2000 (26.05.00)
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 03 August 2000 (03.08.00)
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 03 October 2000 (03.10.00)
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 27 September 2000 (27.09.00)
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 21 September 2000 (21.09.00)
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 30 June 2000 (30.06.00)
 ; NUMBER OF SEQ ID NOS: 15438
 ; SOFTWARE: Molecular Dynamics Sequence Listing Engine
 ; SEQ ID NO 6112
 ; LENGTH: 7997
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL034555.2
 ; FEATURE:
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 12
 ; FEATURE:
 ; OTHER INFORMATION: EST_HUMAN HIT: AU117052.1, EVALUE 0.00e+00
 ; FEATURE:
 ; OTHER INFORMATION: SWISSPROT HIT: P08640, EVALUE 4.00e-10
 ; FEATURE:
 ; OTHER INFORMATION: NT HIT: g11427711, EVALUE 0.00e+00
 ; US-10-203-138A-6112

Alignment Scores:

Pred. No.:	4.03e-10	Length:	7997
Score:	431.50	Matches:	373
Percent Similarity:	34.27%	Conservative:	271
Best Local Similarity:	19.85%	Mismatches:	759
Query Match:	5.34%	Indels:	476
DB:	6	Gaps:	83

US-09-902-432-4 (1-1596) x US-10-203-138A-6112 (1-7997)

Qy 31 ProAlaAlaGluAlaSerGlyAlaAlaGly-AspProAlaAspAlaAspProAla----- 48
 Db 886 CTGGACCAAGAACTCAGGTCTCTCAGACGAGCGCTGCAAAATCTGACTTGTCTAAACT 945
 Qy 49 -----ThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAsnGlyVa 64
 Db 946 GGAATCAGTTAGATGAAGAAAGTACCAAGAAAGGGGCTTTCAAGCCATGTTGAAGTGT 1005

Db 3025 GAATAAAGATTTCAGAACTGAAACTCCACCTTCGGTTGGGCTCCCAAGTGTCAAGTCTCGT 3084
 Qy 647 lGlyGluGluGlnLysProGluGluProLysArgArgValAspThrSerValSerTrpGl 667
 Db 3085 AACTCTAGATACAGCCCATCAGCACTAGAGAAAGCACTGGTGCAAAACCGGTAGAGGC 3144
 Qy 667 uAlaLeuLeuCysValGlySerSerLysLysArgAlaArgLysAlaSerSerAspAs 687
 Db 3145 GCCTTTGGTA-----ACAGAAGAAGAACTGTGGAGCCAGCTACCGCTCTCAGAAGA 3195
 Qy 687 p-----GluGlyGlyProArgThrLeuGlyGlyAspSerHisArgAlaGl 702
 Db 3196 AGCAAGCTGCATCTGAACCTCTCTGCCCCGTG-----GAACAGCTGGA 3243
 Qy 702 uGlu-----AlaSerLysAspLysGluAlaGlyThrAspAlaValPr 716
 Db 3244 ACAAGTAGACCTCCCGCCAGGAGCAGACCCCGATAAAGAGCTGCC-----ATGATGCC 3297
 Qy 716 oAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerProGluPro----- 733
 Db 3298 TGCGGGTGTGAGAA-----GGTTCATCAGGTGACCGCGCTTATCT 3342
 Qy 734 ----AlaGlySerProSerGluGlyGluValSerThrTrpGlu----- 747
 Db 3343 GGATGCCAAGCTTCAACTCCCGGGGCTCGTTTCCCGGAGAGCAACGTAGATCC 3402
 Qy 748 -----SerPheLysArgLeuValThrProArgLysLysSerLysLeuGl 764
 Db 3403 AGAGCTGACAGTACCAGCACTTTCAAAACCAGCTCAGAAAGTCTGAGGAAGCAATGA 3462
 Qy 764 uGluLysAlaGlu-----AspSerSerValGlu----- 773
 Db 3463 GCCAAAGGCGGAAAGCCAGACCCACTGCAGATGCTGAGCTGATGCAAAACCAGAAAGC 3522
 Qy 774 -GlnLeuSerThrGluLeuGluProSerArgGluGluSerTrpValSerLysLysPh 793
 Db 3523 CGAAGCTGCTCTGAGTCTCAGCCCCCAGCTTCTGAAGAT---TTAGAGTTGATCTCTCC 3579
 Qy 793 eileProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGlnAlaThrValGluAs 813
 Db 3580 AGTTGTGCAAGGATAAAAGCAAAACAAAGCAAGCGTTCAAGACCCCTGTTGAG-- 3637
 Qy 813 pSerGlyProValGluLeuAsnGluAspProAsnValProAlaValProLeuSe 833
 Db 3638 -CGAGTGCAGTGAATGATCTGTGAG-----AAGCCCGTCAAGGAAGAG 3681
 Qy 833 rGluTyrAsnAlaValGluArgGluLysMetGluAlaGlnGlnValSerThr----- 851
 Db 3682 TGAG-----AGATAGACCGGAAACTCAAGCGGTCCAATTCCTCGGGAGAGC 3735
 Qy 851 uProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThr----- 867
 Db 3736 ACAGAAGCTTTTGGATTGAAGATGAGGAGAGAGAGATTACAAGGACTGCTTCTAATAAA 3795
 Qy 868 -----LeuValHisThrValSerValAlaValAlaLysGlyThrArgAla-- 882
 Db 3796 CTCTGCTGCAGACCTTGAACATCCCAACCAAGTTTTCCTCTCAGCGCAACAGGCGCG 3855
 Qy 883 ----ValThrSerVal-----GluGluArgSerProSerTrpIl 894
 Db 3856 GAATGAAGAGCGCTTATGCCAACCTATGGGTGACCATGAAACCGCTCTCTCT----- 3907
 Qy 894 eSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProValGl 914
 Db 3908 -----GTCAAAGAGCCCGTTGAGCAA-----CC 3930
 Qy 914 uGluValThrGluLysAspIleleAlaGlu-----GluThrProValLeuThrGlnTh 932
 Db 3931 AAGAGTGACCAAGAAGAGATTGGAGCGAGAGCTTCAGGAGGCTGCAGCGGTCCCAACCAC 3990
 Qy 932 r-----LeuProGluGlyLysAspAlaHisAspAspMetValThrSerGl 947
 Db 3991 CCTCGAGGGGAGGCGCTTCCAAAGACACGCGCGGAGCGGATGAAGAGGAGGAGAACGA 4050

Qy 947 uValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGl 967
 Db 4051 GGCCAGGAACCTGCAGAAACACACTCAAGCCACTGAGGATGCGCGTCCCAAGGTCCCA 4110
 Qy 967 uGluValThrGluAlaSerGlyAlaGlu-----GluThrThrAspMetValSe 983
 Db 4111 GAAAACTGCAGCTGGTGGACCCCAAGGGAAAAAGGAAAAAATGAACCGNAGGTGGA 4170
 Qy 983 rAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGl 1003
 Db 4171 TGCTACA-----CGTCTGAGCCACCACTGAGGTGGGCCCCCAATAGG 4215
 Qy 1003 uValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaLysLeuGl 1023
 Db 4216 CGTGAAGAGAGAGTCTCATGGAACCCCAAGGCTGCTGAGGAGGAGGAGGAGTGAACAGAA 4275
 Qy 1023 nAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnAr 1043
 Db 4276 AGTGACGAAAGATGCTGGCAGACGAAAAACCCCTGAAACCCGCCCTGTTGAA-- 4333
 Qy 1043 gThrGlySerLysAlaLeuGluLysValGluValGluGluAspSer----- 1059
 Db 4334 -----GTGTAGAGAAAAAACCGGCCCTGAAAAAACTCCAAATCAAAGAG 4380
 Qy 1060 -----GluValLeuAlaSerGluLysGluLysAs 1069
 Db 4381 AGGAAGATCTCGAAACTCCAGGTTAGCAGTGACAAATCTGCAAGTCTGAAAAATGTGA 4440
 Qy 1069 p-----ValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGl 1087
 Db 4441 TGCTGTGTCAGTCCCGGGGGCTGCACAGCAGGCGGGAGAG-- 4486
 Qy 1087 ySerGluThrGly-----GlnAlaThrProGluSerLeuGluValProGluValThrAl 1105
 Db 4487 ----GNAATCTGGGTGGTGGCAGTCTCCCTCTGAAAAAGTGAGAGTCCCAAAAGAGGA 4542
 Qy 1105 aAspValAspHisValAlaThrCysGlnValLysLeuGlnGlnLeuMetGluGlnAl 1125
 Db 4543 TGGTTTATCA-----TCCAGTGTGAAAGTGATCC 4572
 Qy 1125 aValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLe 1145
 Db 4573 AGTTGATCAGACAGGAACCCAGAGAAAGACAGACGTTGCTGCTCTGGGCCGTCGCCAGA 4632
 Qy 1145 uAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSer-- 1162
 Db 4633 AGCC---ACCCAGTTAGCCCAAGCAGATGGAGTGGAGCAGGCGGTGGAACACATCCCAA 4689
 Qy 1163 ----GlnAspSerLysAlaThrAlaAlaValArgGlnSer----- 1174
 Db 4690 GCTCGTGAGGCTCTGCTCTGCTTATAGGCAGATGCACAGAGGGCGCTTCCGCC 4749
 Qy 1175 -----GlnValThrGluGluGluAlaAlaThrAlaGlnly 1186
 Db 4750 AGAGGACAGGACAAAGCTTCACACCAAGCAAGTGAACAGAGCTGCTGCGGCCATCGG 4809
 Qy 1186 s-----GluGluProSerThrLeuProAsnAsnValProAlaGlnGluHis-- 1202
 Db 4810 CTCATCATCATGATCTTCTGGGAGCGCAAAACTTCCAGCAGCACTCCACCTTATCC 4869
 Qy 1203 -GlyGluGlu-----ProGlyArgAspValLeuGluProThrGl 1215
 Db 4870 TGGAGAACTCCAGACAGATCTCAACCCCGCAGGTGACAGGCGCTGCAGCTTCTGA 4929
 Qy 1215 nGlnGluLeu-----ThrAlaAlaAlaValProValLeuAlaLysThrGluValGlyGlnGl 1234
 Db 4930 GGAAGGAATGGAGACAGATGAGGCTGTATCTGCGCATCTCGAAACTGAGGCTGTACAGA 4989
 Qy 1234 uGlyGluValAspTrpLeuAspGlyGluLysValLysGluGluGlnGluValPheValHi 1254
 Db 4990 ATCTTCTAGGCGCTCCAGTCAATGCTCTGACCCCTCA----- 5026


```
QY 80 nGluGlu-----GluValValAs 86
|||
Db 4677 AGAGGAGCTGCAGCATCTCGGCAAGAGCTCAGAGGCAGAGATCCAGGCCAAGGCCCCAGCA 4736

QY 86 pGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArgValGluGluMe 106
|||
Db 4737 GGTGGAGGCTGCAGAGCGCAGCGCATTTGAGAAAGAGATCCGCGTAGTCTCGTCT 4796

QY 106 tAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGluGluThrSerG1 126
|||
Db 4797 GCAGCTAGAGACAAC-----GAGCGTCAGCTGGAGGGCGGAGGATGAGCTGCA 4847

QY 126 uIleIle-----GluGlnIleProAlaSerGluAsnValGluGluMetVa 142
|||
Db 4848 GGCTCTCGCTGCACGGGCTCAGGAGGCAGAGCAGACAGAGCGCAGGCTCAGGAGGAAGC 4907

QY 142 lGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValG1 162
|||
Db 4908 CGAGCGCTTCGGAGGCAGGTGAGGATGAGAGCCAGCCAGCAACGGCAGCGGAGGCCGA 4967

QY 162 yPheLysPheThrValLysLysAsp-----LysAsnGluLysSerAspThrValG1 179
|||
Db 4968 GCTGGCCCTCGGTGTAAGCAGACAGAGCGGAGCGAGCGGAGAGAGCAGCGGCCCTGCA 5027

QY 179 nLeuLeuThrValLysLysAspGluGlyGluGlyAlaGlu-----AlaSe 194
|||
Db 5028 GGCTCTCGATGAACCTGAACCTGAGGCGCAGGAGGCGGAGCGGAGCGGCTGTGCCAAGCCGA 5087

QY 194 rValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGly----- 210
|||
Db 5088 GGCAGAGAGGCTGCGCAAGTCAGGTAGCCCTGGAGACAGCGAGCTAGTCAGAAAGT 5147

QY 211 -GluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLe 230
|||
Db 5148 GAGCTGCAGACAGCGCTCGCTCTTTCAGAGAGACCGCACAGTTGAGGCGCAGCT 5207

QY 230 uLysGlnGluGln---SerSerThrGluIleProLeuGlnAlaGlu-----SerAs 246
|||
Db 5208 GCAGGAAGACACGTGACAGTGCACAGCTGCGGAGGAGCGGAGCGGAGCGGAGCGCA 5267

QY 246 pGlnAlaAlaGluGluGluLysAspGluGlyGluGluGluGluGluGluProTh 266
|||
Db 5268 GCAGGCTGAAGCCGAGCGAGCCGCTGAGGAAGCGGAGCGGAGCTGGAGCGCTGGCAGCT 5327

QY 266 rLysSerProGluSerProSerProValAsnSerGluThrThrSerSerPheLysLy 286
|||
Db 5328 GAAGGCCAATGAGCGCTGCGGCTGCGGCTGAGGAGAGAGAGGTGCGACAGCAGAGAG 5387

QY 286 sPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysG1 306
|||
Db 5388 CCTG-----GCCAGCGCGATGCGGAGAAAGCAGGAAGGA 5420

QY 306 uAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluG1 326
|||
Db 5421 AGAGCGCAAGACCGGAAGCCCGGCGGCGGCAAGGCGAGAGCAGCGCCGCTGGCGGAGC 5480

QY 326 uGluLysGluLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspG1 346
|||
Db 5481 AGAGCTGGCTGAGCAGAG---CTGGAGAGAGCGCGCAGCTCAGAGAGGCGCACCGCCCA 5537

QY 346 nAlaArgLeuSerAlaAspTyrrGluLysValGluLeuProLeuGluAspGlnValGlyAs 366
|||
Db 5538 GCAGCGCTGGCTGCCAGCAGAGGCTGATTTCGCTTCGCGGCGAGAGCAGCGGAGCGAGTGA 5597

QY 366 pLeuGluAlaSer---SerGluGluLysCysAlaProLeu-----AlaThrGluVa 382
|||
Db 5598 GCATACGCGGAGCTGTCGGAGGAAGAGCTGGCCGCGCTACAGCAGCAAGCGCAGCAGC 5657

QY 382 lPheAspGluLysMetGluAlaHisGlnGlu-----ValValAlaGluValHisVa 399
|||
Db 5658 CACACAGAGCGCGAGGCTGAGGCTGAGCTGGCGAAGGTTTCGGGCGCAGAGATGAGGT 5717
```

Db 6723 ACCTGAGCGACCGCACTTGAAGCAAGTGAAGAGCGCCGAGCGCTGAAGCAAGTCG-- 6780
 Qy 711 yThrAspAlaValProAlaSerThrGlnGlnGlnAspGlnAlaGlnGlySerSerSerPr 731
 Db 6781 -----GCAAGAGACAGAGCTCAGCGCCAGGCCAGCGGAGCGGCTCG 6824
 Qy 731 ogLupProAlaGlySerProSerGlnGlnGlnValSerThrTPrGlnSerPheLysAr 751
 Db 6825 AGAG----- 6828
 Qy 751 gLeuValThrProArgLysLysSerLysSerLysLeuGlnGlnLysAlaGlnLysPheSe 771
 Db 6829 -----AAACTGCGCAAGAGAGCGGAGCGGAGCGCGCGCTGAGGCGCA 6872
 Qy 771 rValGlnLysLeuSerThrGlnLysLeuProSerArgGlnGlnGlnSerTPrValSerL 791
 Db 6873 GCGCGAGCGAGCGCTTGAACAGAAAGCAGCGGAGCGGAGAGATGAGAAAGCACA 6932
 Qy 791 sLysPheLLeProGlyArgArgLysLysArgAlaAspGlyLysGlnGlnLysAlaThrVa 811
 Db 6933 GAAGTTTGCAGACAGCAACCTACGCGCAAGAGCTCAGGTAGAGCAGAGCTGACACGCT 6992
 Qy 811 lGluAspSerGlyProValGlnLysLeuGlnLysAspAspProAsnValProAlaValAlP 831
 Db 6993 GAGG-----CTGACGCTCGAGAGAGACCGAC--CACCAAGAGCATCTCGA 7037
 Qy 831 oLeuSerGluTyrAsnAlValGlnLysArgLysLysLysLysLysLysLysLysLysL 850
 Db 7038 T---GAGGAGCTGCAGCGCTAAAGCTGAGTAAACAGAGCGAGCGCGGAGCGTAA 7094
 Qy 850 uLeuProGlnLysLeuGlyAlaValTyrValSer---GlnGlnLysSerLysThrLeuVa 869
 Db 7095 GGTAGAGAGAGAGCT 7146
 Qy 869 lHisThrValSerValAlaValLLeAspGlyThrArgAlaValThrSerValGlnGln 889
 Db 7147 -----CTCAAGGCTCGCATGGAACCTGAACCG 7175
 Qy 889 gSerProSerTPrLysSerAlaSerValThrGlnPro---LeuGlnHisThrAlaGln 908
 Db 7176 GCGACCATCTCTCGTGAACAAGCAACACAGCGCTCTCTGAGAGAGAGCGGAGAA 7235
 Qy 908 uAlaMetProProValGlnGlnValThrGlnLysAspLLeLLeAlGlnGlnThrProVa 928
 Db 7236 GATGAACAAGGTGGCAGAGAGAGCTGACGCTGAGCTGAGCTGAGAGAGAGAGAG 7295
 Qy 928 lLeuThrGln-----ThrLeuProGlnGlnLysAspAlaHisAspAspMetVa 944
 Db 7296 GCTGCGGAGCTTACCGGAGAGAGAGCTGCGGAGCGGAGCGGAGCGGAGAGATGCT 7355
 Qy 944 lThrSerGlnValAspPheThrSerGlnAlaValThrAlaThrGlnThrSerGlnAla 964
 Db 7356 GAAGGAAAGATCCAGCGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7415
 Qy 964 uArgThrGlnGlnValThrGlnAlaSerGlyAlaGlnGlnThrThrAspMetValSerAl 984
 Db 7416 GCAGCAGCAGAGAGAGAGCTG-----GCACAGAGAGAGAGAGCGCGGCTGACAG 7463
 Qy 984 aValSerGlnLysThrAspSerProAspThrThrGlnGlnAlaThrProValGlnGln 1004
 Db 7464 G-----GACAAAGAGCAAAATGCTTCACAGCTGAGTAAAGGA 7499
 Qy 1004 lGluSerGlyValLeuAspThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1024
 Db 7500 GACACAGGCTTCCAGCGAGCCTGAGAGGCTGAGCGGAGCGGAGCGGAGAGAGAG 7559
 Qy 1024 aValAlaAspLysValLys----- 1030
 Db 7560 AAGAGCTGAAGCGCTCAAGTTGCGCATGCTGATGATGATGATGATGATGATGATGAT 7619
 Qy 1031 -GlnGlnSerGln----- 1034

Db 7620 GGAGGATGCGCCAGCGCTTCCGAGACAGAGCTGAAGATTCGCGCAAAAGCTCCACCG 7679
 Qy 1035 -----ValProAlaThrGlnThr-----ValGlnArgThrGln 1045
 Db 7680 TGAATCTCTACACAGAGAAAGTGATGATGATGATGATGATGATGATGATGATGAT 7739
 Qy 1045 ySer-----LysAlaLeuGlnLysValGlnGlnValGlnGlnGlnGlnGlnGlnGln 1060
 Db 7740 GAGTACCAAGATGCCGAGCGCTGAGGAGCGCCATTCCTGAGCTGAGCGGAGAGGA 7799
 Qy 1060 u-----ValLeuAlaSerGlnLysGlnLysAspValMe 1071
 Db 7800 GAAGCTCAAGCAGAGAGGAGGAGTACTGACGTCAAGTCAAGTCAAGTCAAGTCAAGT 7859
 Qy 1071 rProLysGlyProValGlnGlnAlaGlyAlaGlnHisLeuAlaGlnGlnGlnGlnGln 1091
 Db 7860 GCAGAGCAGATACGTGAGAGAGACACAGCGCTGAGAGAGAGCTTCTCTGAGAAAG 7919
 Qy 1091 yGlnAlaThrProGlnSerLeuGlnValProGlnValThrAlaAspValAspHisValAl 1111
 Db 7920 CAGCTTGTCAACCGCAACGCTTCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7979
 Qy 1111 aThrCysGlnValLLeLysLeuGlnGlnLysLysLysLysLysLysLysLysLysL 1125
 Db 7980 CCAGAGACAGAGTGGCAAAAGCAAAACAGCTGACAGAGAGAGAGAGAGAGAGAGAG 8039
 Qy 1125 aValAlaProGlnSerSerGlnThrLeuThrAspSerGlnThrAsnGlySerThrPro 1145
 Db 8040 GATGAGCAGAGAAAGAGAGAGAGCTGTGAGCAGCATGAG--GAGGCCGAG 8090
 Qy 1145 uAlaAspSerAspThrAlaAspGlyThr----- 1154
 Db 8091 GCGGAGCGGTGAGGAGAGAGAGGTGTGAGGAGGAGAGAGAGAGAGAGAGAGAGAG 8150
 Qy 1155 -----GlnGlnAspGlnThrLLeAspSerGlnAspSerLysAlaThrAlaAlaVa 1171
 Db 8151 GCAGAGCGGAGCAGCAGAGAGAGAACTACTGACAGAGAGAGAGAGAGAGAGAGAG 8210
 Qy 1171 lArgGlnSerGlnValThrGlnGlnAlaAlaThrAlaGlnLysGlnGlnProSerTh 1191
 Db 8211 GCTGAGCGCTGAG 8270
 Qy 1191 r-----LeuProAsnAsnValProAlaGlnGlnGlnHisGln 1203
 Db 8271 CTCCAGAGCTGCTGCCACAAAGCACTGCCAAT----- 8304
 Qy 1203 yGlnGlnProGlyArgAspValLeuGlnProThrGlnGlnGlnLysLeuThrAlaAla 1223
 Db 8305 -----GCCCGAGCGCATGATGCGCCCTCCATGAG-- 8337
 Qy 1223 lProValLeuAlaLysThrGlnValGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1243
 Db 8338 -----GCCAGCGGAGTACACCTTGAGGAGATTAAGTCAAGAGTCCAGCTCA 8387
 Qy 1243 uLysValLysGln-----GlnGlnGlnValPheValHisSerGlyProAsnSe 1259
 Db 8388 GCAGCTACAGAGAGAGAGAGATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 8447
 Qy 1259 rGlnLysAlaAlaAspValThrTyrAspSerGlnVal-----Me 1272
 Db 8448 CACCAAGGTGCTGAGCTCAGCAGCAGAGAGATGAGAGAGAGAGAGAGAGAGAGAG 8507
 Qy 1272 rGlnValAlaGlyCysGlnGlnLysGlnSerThrGlnValGlnSer 1287
 Db 8508 CAGATCGCAGAGATTCCTCTTAAGCCACCAATAGAACTGAGTGTACACAGCCCT 8567
 Qy 1288 -----LeuSerLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1297
 Db 8568 ACAGCGGAGCTCTCAGCGCTGAAACAGCCCTTATCTTACTTGAAGCCGAGAGAG 8627
 Qy 1297 r-----AspValGlnLysGlnLysArgGlnThrLysLysProGlnGlnValSerGln 1313
 Db 8628 GGGCTTCTGCTGAGACCTCTGCGAGACCGGCGCTGACGCTCAATGAGCGCTGAAAG 8687

```
QY 1313 u----- 1313
DB 8688 GGGTGTGGTCCGAGCTGCACCACAAAGCTGCTGCAGCTGAGCGTCCGCTCACTGG 8747
QY 1314 -----GluGlyGluGlnGluThrAlaAlaProGluHisGlu----- 1325
DB 8748 CTACAAGGACCTTACACAGGAGACAGAT-CTCTCTCTTCACAGGCGATGAAGAAGACC 8806
QY 1326 -----GlyThrTyrglyysProValLeuThrLeuAsp----- 1336
DB 8807 TCATTGTCAGGACCATGGATCGCTGCTGGAAGCCAGATGCCACAGGTGGCATCA 8866
QY 1337 -----MetProSerSerGluArgGlyLysAlaLeuGlySerLeu----- 1349
DB 8867 TTGACCTGTACACAGCACCGGTGTTCCGCTGAGCGGTCTACACAGCGGTACTTCG 8926
QY 1350 -----GlyGlySerProSer-Leu-----ProAspGlnAspLysAlaGlyCysIleGluV 1366
DB 8927 ATGAGGAGATGAACCGGTGCTGGCTGACCCAGCGATGACACCAAGGGCTCTTTGACC 8986
QY 1366 alGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValI 1386
DB 8987 CCAACTCTAC-----GAGAAGCTCA 9007
QY 1386 leGluThrValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuL 1406
DB 9008 CGTACCTGTCAGCTGCTGGAGCGGTGTGTGGAGGACCCCGAG---ACAGGCTCGGCTCC 9064
QY 1406 euProAlaGluLysSerSerAlaThrGlyHis-Trip---ThrIeuGlnHisAlaGlu 1424
DB 9065 TGGCACTCACAGACAGGTCGCAAGGGTGTGTGAGTGTGTACATGACACCGAGGCC 9124
QY 1425 AspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleValThrPro 1444
DB 9125 GTGACGTCTCGAAAGGCCACAGTGTCTGCACCATTCGCGCATGTCAGGGCAAGACC 9184
QY 1445 AlaProGluSerThrLeuHisProAspLeuGlnGlyLysSerAlaSerGlnArgGlu 1464
DB 9185 TGACCATCTGGGAGA-----TCATCACTCAGAGTACT 9217
QY 1465 ArgSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThr 1484
DB 9218 TCACAGCGGAGCAGCAGCGGACCTGCTCCGCGCATGTCGCCAGC-GGCCGC-----ATC 9270
QY 1485 AlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsn 1504
DB 9271 ACGGTGGAGAGATCATCAGATTGTCATCAGCGTGTAGAGGAACACGAGCGGAAGGCC 9330
QY 1505 LysIle-ValLeuAsnValIleGlnThrAlaValAspGlnPheAlaAargThrGluThrAl 1524
DB 9331 CAGCTCTGCTTTG----- 9343
QY 1524 aProGluThrHisAlaTyraSpSerGlnThrGlnValProAlaCysArgLeuAspSerAr 1544
DB 9344 -----AGGCGCTCCGTCGCTTGTGCTGTGTCAGAGC-- 9376
QY 1544 gGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLys 1559
DB 9377 -----TGCTGGACA---GTGGAGTCATCATGTATGAAG 9406

RESULT 10
US-09-724-676-25032
; Sequence 25032, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25032
```

```
; LENGTH: 14335
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-25032

Alignment Scores:
Pred. No.: 2,38e-09 Length: 14335
Score: 418.00 Matches: 334
Percent Similarity: 35.65% Conservative: 255
Best Local Similarity: 20.22% Mismatches: 701
Query Match: 5,18% Indels: 364
DB: Gaps: 61

US-09-902-432-4 (1-1596) x US-09-724-676-25032 (1-14335)
QY 5 SerSerThrGluGlnArgSerProGlu-GlnProAlaGlySerAspThrProSerGluLe 24
DB 4187 AGCAGCAGCGGCGAGAGAGCGGCGCGCTGGCCGAGGTGGAGGCGCGCTCGAGAAGC 4246
QY 24 uValLeuSerGlyHisGly-----ProAlaAlaGluAlaSerGlyAlaAlaGlyAspPr 42
DB 4247 AGCGGCAGCTGGCGGAGCGCACGCCCAAGGCACAGCGGAGCGGAGCGGAGGCGAAG 4306
QY 42 oAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAs 62
DB 4307 AGCTGCAGCAGCGCATGCA-----CGAGGAGGTGCTGCGCGGAGGAGCGCG 4354
QY 62 nGly-ValAlaGluGlnGlyAspValHisValGlnGluGlnGluGlnGluGlnGluG 82
DB 4355 CGGTGAGCGCGCAGCAGCAGAGCGCAGCATTCAGAGGAGGTGCAGCAGCTCGCGCAGA 4414
QY 82 luGluValValAspGluAspValGlyGlnArgLysGlu----- 95
DB 4415 GCTCGGAGCGGAGATCCAGGCCCAAGGCCGAGCAGAGCGGCGGTGAGCGAGCGGC 4474
QY 96 --AspValArgGluLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluA 115
DB 4475 TGCATCATCAGAGGAGATCCCGTGTGCTGCTGTCAGTTGGAGGCCACCGAGCGCAGC 4534
QY 115 spileThrLysAspGlyGlnGluGluThrSerGluIle---lleGluGlnIleProAlaAs 134
DB 4535 GTGCGCGGCTGAGGGGAGCTGCGGCACTGCTGTCACGGCGGAGGAGGTGAGGCAC 4594
QY 134 erGluAsnAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyP 154
DB 4595 AAAAGCGACAGCGCAGGAGGAGCGCGCTGCGGAGGAGGTGTCAGGACGAGCGC 4654
QY 154 helLysLysValPheLysPheValGlyPheLysPheThrVallLysLysAsp-----L 171
DB 4655 AGCGTAAGCGGAGCGGCGGAGGTGAGCTGCGCTGCGCGCTGAAGCGCGAGACCGAGCGG 4714
QY 171 ysAsnGluLysSerAspThrValGlnLeuThrVallLysLysAspGluGlyGluGlyA 191
DB 4715 CGCGCGAGAGAGCAGCGGCGCTGACAGGCTGAGAGAGTGTGCGGTGCGGCGGAGGAG 4774
QY 191 laGlu-----AlaSerValGlyAlaGlyAspHisGlnGluProSerValG 206
DB 4775 CGGAGCGGCGCTGCGGAGCGCGAGGTGAGCGGCGGAGGTACAGGTGCGCTG 4834
QY 206 luThr-----AlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnS 222
DB 4835 AGACGCGCAGCGCGCTGTCAGAGCGGAGTGCAGAGCAAAAGCGCGCTCTTCTCGCGCAGA 4894
QY 222 erThrLysLysGlnGluGlyThrLeuLysGlnGln---SerSerThrGluIleProL 241
DB 4895 AGACGCGCAGCTGGAGCGCTCCCTGAGGAGGAACACCTGCTGTGGCACACCTCGCGG 4954
QY 241 euGlnAlaGlu-----SerAspGlnAlaAlaGluGluAlaLysAspGluGlyG 258
DB 4955 AGGAGCTGAGCGGCGGCGGACAGCAGCAGCGCGCGGCGGCGGCGGCGGAGGCGCAG 5014
QY 258 luGluLysGlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnS 278
```

Db 5015 AGCGGAGCTGAGCGCTGACAGCTCAAGCCACAGAGCGCTACGCGCTGACAG 5074
 Qy 278 eRgluThrThrserserPheLysPhePheThrHisgluTyrAlaGluTyrAlaGlu 298
 Db 5075 CGAGGAGGCTGCGCGACGAGAGAGCTG-----GCCG 5107
 Qy 298 yRThrserserHisLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGlu 318
 Db 5108 AGCGCGAGGCTGAGGACGAGAGAGAGAGCGGAGCGCAGCGCGCGCGCGAGAG 5167
 Qy 318 IngluAlaGluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu 338
 Db 5168 CGAGAGAGAGAGCGCTCGCGCGAGCGGAGCTGCTGACAGAG-----CTGAGAGAGAG 5224
 Qy 338 IngluProAlaGluAspThrAspGluAlaArgLysSerAlaAspTyrGluLysValGlu 358
 Db 5225 GCGAGCTGCGGAGAGGACACCGCGCAGACAGCGCTGCGCGCGGAGCGAGTTCGCG 5284
 Qy 358 euProLeuGluAspGluValGluAspLeuGluAlaSer---SerGluGluLysCysAlaP 377
 Db 5285 TCAGGCGCCGAG 5344
 Qy 377 roLeu-----AlaThrGluValPheAspGlu----- 385
 Db 5345 GCGCTGAGGCTGAGCGCGCTGCGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 5404
 Qy 386 --LysMetGluAlaHisGluGluVal-----AlaGluValH 398
 Db 5405 CCAAGGCTGCGGCGCGAGATGAGTGTCTGCGCAGAGCGAGCGAGCGCTGAGAGAG 5464
 Qy 398 lValSerThrValGluLysThrGluGluGluGluGluGluGluGluGluGluGlu 418
 Db 5465 CCGCGCTCCACGAGCGAGAGTCCAGACAGAGCTG-----GAGCGAGAGCGCG 5512
 Qy 418 lValValAlaGluGluThrGluGlu-----SerLeuProProGlu---- 431
 Db 5513 GCGCGTTCGCGAGCTGCGCGAGAGAGCGCGCGCGCGCGCGCGCGAGAGAGAG 5572
 Qy 432 -----LysLeuAlaGluProGluGluValProGluGluAlaGluProAlaGlu 448
 Db 5573 AGCGGAGCGGAGCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5629
 Qy 448 lLeuMetLysSerArgLysMetCysValSerLysLysAspHisThrGluLeuThrAsp 468
 Db 5630 GGGTCTTCGAGAGAGCTGCGCGCATCGC-----GAGCGCATCGCGC 5674
 Qy 468 euSerProGluGluLysThrLeuProLysHisProGluGluLysValSerGluValGlu 488
 Db 5675 TCAAGACGAGAGCGGAGATCGCGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 5734
 Qy 488 euLeuSerSerGluGlu-----ArgLysValGluGluLysSerProLeu 503
 Db 5735 GCGTGCAGAGAGCGAGCGCTTCACAGCGCGCGCGCTGAGAGAGAGAGAGAGAG 5794
 Qy 503 yAluLysLeuHeSerSer-----SerGluLysLysLysSerGlu-----Lys 518
 Db 5795 AGCGTGCATCGAGAGAGCGCTGCGCATCGCTGCGAGAGAGAGAGAGAGAGAGAG 5854
 Qy 518 yGluLysGluLysArgGluGlu-----GluGluAspGluGluP 531
 Db 5855 GCGAGAGAGAGGCT--GGTGAAGAGACAGCTGAGAGAGAGAGAGAGAGAGAGATC 5913
 Qy 531 roGluGluLysThrHisLysHisThrGluSerProGluSerAlaAspGluGlu-----L 549
 Db 5914 CTGCG-----GCTGAAGAGAGAGCTTCGAGAGAGAGAGAGAGAGAGAGAGAG 5961
 Qy 549 yGluLysSerSerAlaSerSerProGluGlu--ProGluGluThr---ThrCysLeuGlu 567
 Db 5962 CTGAGACTGAGAGCTGAGAGCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6021
 Qy 568 LysGluProLeuGluAlaProGluAsp----- 576
 Db 6022 CAGGCGAGCTGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6081

Qy 577 GlyAlaGluGluGluThrThrSerAsp-----GlyGluLys 589
 Db 6082 CGTGAGCTGAGAGAGCGCTGCAGAGAGCTTCGCGCGAGAGAGAGAGAGAGAGAG 6141
 Qy 590 LysArgGluGluLysThrProThrAlaSerPheLysLysMetValThrProLysArg 609
 Db 6142 CGAGAGCGCGCGCTGAG 6201
 Qy 610 ValArgArgProSerGluSerAspLysGluGluGluGluLysValLysSerAlaThr 629
 Db 6202 CTGCGAG 6261
 Qy 630 LeuSerSerThrAspSerThrValSerLysMetGluAspGluValLysThrValGlu 649
 Db 6262 CAGAGAGAGCTG-----CAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6306
 Qy 650 GluGluLysProGluGluProLysArgArgValAspThrSerValSerTrpGluAla 669
 Db 6307 CAGCAG 6360
 Qy 670 lLysValGlySerSerLysArgAlaArgLysAlaSerSerSerAspAspGlu 689
 Db 6361 GACCACTGCGCGAG 6420
 Qy 690 -----GlyProArgThrLeuGluLysAspSerHisArgAlaGluGluAlaSer 706
 Db 6421 CCGGTGAGCGGAG 6480
 Qy 707 AspLysGluAlaGluThrAspAlaValProAlaSerThrGluGluGluAlaGlu 726
 Db 6481 CTGAGAGAGCTG-----GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6522
 Qy 727 GlySerSerSerProGluProAlaGlySerProSerGluGluGluValSerThrTrp 746
 Db 6523 GCACAGCGCGCTGAGAG----- 6540
 Qy 747 GluSerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGlu 766
 Db 6541 -----AAAGCTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6570
 Qy 767 AlaGluAspSerSerValGluGluLeuSerThrGluLysProSerArgGluLys 786
 Db 6571 GCGCGCGCGGAG 6630
 Qy 787 TrpValSerLysLysPheLysProGluArgArgLysLysArgAlaAspLysGlu 806
 Db 6631 ATGAGAGAGCATTAAGAAATTCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6690
 Qy 807 GluGluAlaThrValGluAspSerLysProValGluLeuAspGluAspProAsnAl 826
 Db 6691 GAGCTGACAACTGCG-----CTGACGCTGAGAGAGAGAGAGAGAGAGAGAG 6729
 Qy 827 ProAlaValProLeuSerGluTyrAsnAlaValGluArgGlu-----Lys 842
 Db 6730 -----CACAG 6771
 Qy 843 MetGlu-----AlaGluLysAsnThrGluLeuProGluLeuGluAlaVal 858
 Db 6772 GCGAG 6831
 Qy 859 TyrValSer---GluGluLeuSerLysThrLeuValHisThrValSerValAlaVal 877
 Db 6832 CCGGTGAG 6858
 Qy 878 AspGluThrArgAlaValThrSerValGluGluArgSerProSerTrpLysSerAla 897
 Db 6859 -----CTCAAG 6912
 Qy 898 ValThrGluProLeuGluHisThrAlaGluAlaMetProProVal---GluGluVal 916
 Db 6913 AATTAAGAGAGCTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6972

QY 917 ThrGluLysAspIleIleAlaGluThrProValLeuThrGln-----Thr 932
 Db 6973 GCGCGGTGAGTGTGGCGGCGCAGGCTGCGCAGTCTGGCGGAGGAGGAC 7032
 QY 933 LeuProGluGlyLysAspAlaHisAspMetValThrSerGluValAspPheThrSer 952
 Db 7033 CTGGCAGCAGCGCGGCTTGGCAGAGAAGATGCTCAAGAGAAGATGACGGCGGTCCAG 7092
 QY 953 GluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAla 972
 Db 7093 GAGGCCACGCGGACTCAAGGCTGAGCGGAACTGCTGCAGCAGCAGAGAAGAGCTTGCAG 7152
 QY 973 SerGlyAlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerPro 992
 Db 7153 GAGCAGCGCGGCGGTGCGAGGAG-----7176
 QY 993 AspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGlu 1012
 Db 7177 GACAGGAGCAGATGGCGCAGCAGCTGGCGGAGGAGAGCGAGGCTTCCAGCGGACGCTG 7236
 QY 1013 GluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLys-----1030
 Db 7237 GAGCGCAGCGCGCAGCGCAGCTGAGATGAGCGCTGAGCGCTGAGCGCCTCAAGCTCGT 7296
 QY 1031 -----GluGluSerGlnValProAlaThrGlnThrValGlnArgThrGlySer 1046
 Db 7297 GTGGCCGAGATGAGCGCAGCGCCAGCGCGCTGAGGAGCAGCGCCAGCGCTTCCGGAAG 7356
 QY 1047 LysAlaLeuGluLysValGluGluValGluLysAspSerGluValLeuAlaSerGluLys 1066
 Db 7357 CAGCGGAGGAGATCGGTGAGAGCTGCACCGCAGGAG-----CTCGCCACCCAGGAG 7410
 QY 1067 GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGln 1086
 Db 7411 AAGGTGACCTGTGTGACAGACTGAGATCCAGCAGCAGCAGAGTGACCATGATGCCGAG 7470
 QY 1087 GlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrAlaAsp 1106
 Db 7471 -----CGCCTGCGGAGGCGCATCGCTGAG 7494
 QY 1107 ValAspHis-----ValAlaThrCysGlnValIleLysLeuGlnLeu--MetGlu 1123
 Db 7495 CTGGAGGCTGAGAGGAGAGAGCTCCACAGAGGCGCAACTGTGCGAGCTCAAGTCTGAG 7554
 QY 1124 GlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThr 1143
 Db 7555 GAGATGACAGCGGTGCAGCAGGAGCAGCTGCTCAGCAGCAGCAGCGGCGCTGCAGCAAGC 7614
 QY 1144 ProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGln 1163
 Db 7615 TTCTCTCTGAAAAGGACAGCCTG-----CTACAGCGGAGCGGCTTTCATCGAGCAGGAG 7668
 QY 1164 AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThr 1183
 Db 7669 AAGGCCAAGCTG-----GAGCAGCTTTCAGGAGCAGAGGTGGCCAAG 7710
 QY 1184 AlaGlnLysGluGluProSerThrLeuProAsnValProAlaGlnGluGluHisGly 1203
 Db 7711 GCACAGCAG-----CTGCGTGAGGAG-----7731
 QY 1204 GluGluProGlyArgAspValLeuGluProThrGlnGlnLeuThrAlaAlaVal 1223
 Db 7732 CAGCAGCGGCGCAGCAGCAGATGAGCAGCAAGCGGCGGCTGTTGGTGGCAGCATGGAG 7791
 QY 1224 ProValLeuAlaLysThrGluValGlyGlnGluGly-----GluVal 1237
 Db 7792 GAGCGCGCGCGGCGAGCATGAGCGCGGAGGCGGCTGCGCGCCAGCAGGAGGAGGAGCTG 7851
 QY 1238 AspTrpLeuAspGlyGluLysValLysGluGluGlnGluValPheValHisSerGlyPro 1257
 Db 7852 CAGCAGCTGAGCAGCAGCGG-----CGCAGCAGCAGGAGAGTGTGCTGAGGAGAACCCAG 7908
 QY 1258 AsnSerGlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCys 1277

Db 7909 AGCTCGGTGAGCAGCTGCGAGCTCTGGAGGAGCAGCACCGGCGCGCTGGCGCACTCA 7968
 QY 1278 GlnGluLysGluSerThrGluValGlnSerLeuSer-----1289
 Db 7969 GAGAGGTCACTGCTCGCAGGTGGTGGCCACAAAGACCCTGCCCAATGGCCGGGATGCA 8028
 QY 1290 -----LeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGlu 1305
 Db 8029 CTTGATGGCCCGCGCAGAGGCGGAGCGGAGCAGCATTCGATGGCTCGCGCGGAAG 8088
 QY 1306 ThrLysProGluGlnValSerGluGluGlyGluGlnGluThr-----1319
 Db 8089 GTCTCAGCTCAGAGGTGCGAGGAGCGCGCAT--CTGAGTGGCGGAGAGCTGCGAGCGGTT 8147
 QY 1320 ---AlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetPro 1338
 Db 8148 GCGCGAGGCGCACACCGTGA-----8171
 QY 1339 SerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGln 1358
 Db 8172 -----CGAGCTCGCAGCGGCGGAGAGCGTGGCCACTACTCGAGGCGCGCAGCAG 8222
 QY 1359 AspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThr 1378
 Db 8223 TATCGCAGGCGTGTCTGTAAGGCCAC-----CAATGAGAA 8258
 QY 1379 AlaGluAlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerPro 1398
 Db 8259 GCTGAGTGTTTAGCCGCCCTGCAGAGGCGAGCTGTGAGTCCCGGCACGCGCCCTCATCCT 8318
 QY 1399 Glu-----1399
 Db 8319 GCTGGAGGCGCAGCGCGCTCAGGCTTCTGCTGGAGCCCTGTGCGGAACCGCGGCTGAC 8378
 QY 1400 -----CysValGlyAlaHisLeuLeuProAlaGluLysSer 1411
 Db 8379 CGTCAACGAGGCTGTGAAGAGGCTGTGGTGGCGCCGAGCTGCACCAAGCTGTCTGTC 8438
 QY 1412 SerAlaThrGlyGlyHisThrThrLeuGlnHisAlaGluAspThrValProLeuGlyPro 1431
 Db 8439 GCGCAGCGCGCGCTCAGCTAGCTAGCAAGACCCCTACATGGCCAGCAGATCTCTCTTT 8498
 QY 1432 GluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeuHis 1451
 Db 8499 CCAAGCCATGCAGAGGG-----CCTCATCGTCCGGGAGCAGCGCAT 8540
 QY 1452 ProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluAspLys 1471
 Db 8541 CCG-----8543
 QY 1472 ProAspAlaGlyProAspAlaAspGlyLys 1481
 Db 8544 CTGCTGGAGGCGCCAGATCGCCACGCGCGG 8573

RESULT 11
 US-09-724-676A-25032
 ; Sequence 25032, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724, 676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 25032
 ; LENGTH: 14335
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676A-25032

Alignment Scores:

Pred. No.: 2.38e-09 Length: 14335
 Score: 418.00 Matches: 334
 Percent Similarity: 35.65% Conservative: 255
 Best Local Similarity: 20.22% Mismatches: 701
 Query Match: 5.18% Indels: 364
 DB: 5 Gaps: 61

US-09-902-432-4 (1-1596) x US-09-724-676A-25032 (1-14335)

Oy SerSerThnglIngluYsAspProglu-GlnProAlaglySerAspThrProSerGlulue 24
 Db 4187 AGCAGCAGCGGCGCAGAGAGCGCGCGCTGCGAGGTGAGAGCGCGCGCGAGAAAC 4246
 Oy 24 uValLeuSerGlyHsGly-----ProAlaAlaGlulAsSerGlyAlaAlaGlyAspPr 42
 Db 4247 AGCGGAGCTGCGCGAGCGCGCAGCGCAAGGCAAGCGCGAGCGCGAGCGCGAGAG 4306
 Oy 42 oAlaAspAlaAspProAlaThrLysLeuProGlulnysAsnGlyGlnLeuSerValAs 62
 Db 4307 AGCTGAGAGCGCATGCA-----GAGAGAGGTGTCGCGCGGAGAGAGCGG 4354
 Oy 62 nGly-ValAlaGlulnGlyAspValHsIvalGlnGluLysnGlnGlnGlnGlnGln 82
 Db 4355 CGGTGAGCGCGCAGCAGCAGAGCGCGCATTCAGAGAGAGCTGCGAGCGCGCAGAG 4414
 Oy 82 lGluValValAspGluAspValGlyGlnArgGluSerGlu----- 95
 Db 4415 GCTCGAGAGCGAGATCCAGCGCAAGCGCGGCGAGCAGAGCGCGCTGAGCGCAGCGCGC 4474
 Oy 96 --AspValAlaArgGluLysAspArgValAlaGlnLumetalAlaAlaAsnSerThrAlaValGlu 115
 Db 4475 TCGGATTCAGAGAGAGATCCGCTGTGTCGCTGCGCATTTGAGAGCGCAGCGCGCAGC 4534
 Oy 115 spLeuThrLysAspGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 134
 Db 4535 GTCGCGCGCGCTGAGGCGGAGCTGCGAGCACTGCTGCGAGCGCGGAGAGCGCGAGCGC 4594
 Oy 134 eArgLysnAsnValGlnGlnLumetalGlnProAlaGluSerGlnAlaAsnAspValGlyP 154
 Db 4595 AAAAGCGAGAGCGCGCAGAGAGCGCGCGCTTGCAGAGCGAGTGCAGAGCGAGAGCC 4654
 Oy 154 hElyLysValAlaHelysPheValGlyPheLysPheThrValLysLysAsp-----L 171
 Db 4655 AGCGTAAAGCGGAGCGGAGGTGAGCTGCGCTCGCGCTGAAAGCGCGAGCGAGCGG 4714
 Oy 171 yAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGlnGlnGlnGln 191
 Db 4715 CGCGGAGAGAGAGCGCGCGCTGCGAGAGCTGCGCGCTGCGAGCGCGAGAGAG 4774
 Oy 191 lAgLu-----AlaSerValGlyAlaGlyAspHisGlnGlnProSerValG 206
 Db 4775 CGAGAGCGCGCTGCGCGAGCGCGAGTGAAGCGCGGCGAGTGAAGTGGCCCTG 4834
 Oy 206 lInThr-----AlaValGlyGlnSerAlaSerLysGlnSerGlnLumLysGlns 222
 Db 4835 AACCGCGCAGCGCAGCTGAGAGCGGAGCTGCGAGCAAAAGCGCTCTTCGCGAGAA 4894
 Oy 222 eThrGluLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 241
 Db 4895 AAGCGCAGAGCTGAGCGCTCTCTGAGAGAGAACACGTGCTGCGAGCGAGCTGCGG 4954
 Oy 241 euGlnAlaGlu-----SerAspGlnAlaAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 258
 Db 4955 AGAGAGCTGAGCGCGGCGCAGCAGCAGCGCGAGCGCGAGCGCGCGCGAGAGAGAG 5014
 Oy 258 lGluLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 278
 Db 5015 AACGGAGCTGAGCGCTGCGAGCTCAAGCGCAACGAGCGCTTACGCGCTGCGAGAG 5074
 Oy 278 eArgLysThrSerSerPheLysPhePheThrHisGlyTrpAlaGlyTrpArgLysL 298
 Db 5075 CGAGAGAGGTGCGCGCAGAGAGAGCGCTG-----GCCG 5107

Oy 298 yThrSerPheLysLysSerLysGlnLysAspLeuGlnThrAlaGlnLysArgLysGln 318
 Db 5108 AGCGCGAGCGCTGAGAGAGAGAGAGAGCGGAGCGCGAGCGCGCGCGCGAGAG 5167
 Oy 318 lGlnAlaGlnLysValAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 338
 Db 5168 CGAGAGAGAGCGCGCTCGCGAGCGGAGCTGCGTGAACAAGAG-----CTGAGAGAGAGC 5224
 Oy 338 lGlnProAlaGlnLysAspThrAspGlnAlaArgLeuSerAlaAspTrpGlnLysValGln 358
 Db 5225 GGCAGCTGCGGAGAGAGCGCGCGCAGCGAGCGCGCTGCGCGAGCAGAGAGGTTCGCGC 5284
 Oy 358 euProLeuGlnAspGlnValAlaLysLeuGlnLysAsp-----SerGlnGlnLysCysAlaP 377
 Db 5285 TCGGCGCGAG 5344
 Oy 377 rOlEu-----AlaThrGlnValPheAspGlu----- 385
 Db 5345 GGTGCGAGCTGAGCGCGCTGCGAGCGCGCAGCAAAAGCGCAGAGCTGAGAGCGAGCTG 5404
 Oy 386 --LysMetGlnAlaHisGlnGlnValVal-----AlaGluValH 398
 Db 5405 CCAAGTTCGCGGCGAGATGAGAGGTGCTGCGCAGCAGAGCGAGCGAGCTGAGAGAGAGT 5464
 Oy 398 lSValSerThrValGlnLysThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 418
 Db 5465 CGGCTTCACAGCAGAGAGAGTCAAGCAGAGAGCTG-----GAGCGCGAGCGCG 5512
 Oy 418 lYValValValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 431
 Db 5513 GCGGTCGCGAGAGCTGCGCGAGAGAGCGCGCGCGCGCTGCGCTGCGCGAGAGAGAGCA 5572
 Oy 432 -----LysLeuAlaGluProGlnGlnValProGlnGlnAlaGluProAlaGln 448
 Db 5573 AGCGCAGCGCGCAGCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5629
 Oy 448 lLumMetLysSerArgGlnGlnMetCysValSerGlyGlyAspHisThrLumLeuThrAsp 468
 Db 5630 GGGTCTTGCAGAGAGAGTGGCGCGCGCATCGGC-----GAGCGCAGCGCGC 5674
 Oy 468 euSerProGlnGlnLysThrLeuProLysHisProGlnGlnGlnGlnGlnGlnGlnGln 488
 Db 5675 TCAAGCAGAGCGGAGAGATCCCGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5734
 Oy 488 eLysSerSerGlnGln-----ArgLysValGlnGlnGlnGlnGlnGlnGlnGlnGln 503
 Db 5735 GCGTGCAGAGAGAGAGAGCTTCCAGCGCGCGCGCTGAGAGAGAGAGAGAGAGAGAG 5794
 Oy 503 yLysLeuPheSerSer-----SerGlyLeuLysLysLeuSerGly-----LysL 518
 Db 5795 AGGCTGACATCGAGAGAGCGCTGCGCGCAGCTGCGAGAGAGAGAGAGAGAGAGAG 5854
 Oy 518 ySGLnLysGlyLysArgLysGly-----GlyGlyAspGlnGlnGlnGlnGlnGlnGln 531
 Db 5855 GGCAGAGAGAGAGCTGATC 5913
 Oy 531 rGlnGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 549
 Db 5914 CTGCG-----GCTGAAGGCGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5961
 Oy 549 ySGLnLysSerLysSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 567
 Db 5962 CTGAGAGCTGAG 6021
 Oy 568 LysGlyProLeuGlnAlaProGlnAsp----- 576
 Db 6022 CAGCGCAGAGCTGAG 6081
 Oy 577 GlyGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 589
 Db 6082 CGTGAAGCTGAG 6141
 Oy 590 LysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 609

Db	6142	CGGAAGCGCGCTGGAGGAGTTCGAGCGGCTGAAAGCAACGTCGAGGAGCGCGCGCGC	6201
Qy	610	ValArgArgProSerAspIysGluGluGluLeuGluLysValLysSerAlaThr	629
Db	6202	CTGCGGAGCGACGGAGCAGGAGTCCGCGCGCAGCTGCGAGCTGGCCAGGAGCGCGC	6261
Qy	630	LeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGlu	649
Db	6262	CAGAAGCGGCTG-----CAGGCGGAAGAGAAGACACGCCCTTCGCGGTG	6306
Qy	650	GluGlnIysProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeu	669
Db	6307	CAGCAGAAGGAGCAGAGCTACACAGACGCTGCAGCAGGACGACAGC-----GTGCTG	6360
Qy	670	IleCysValGlySerSerLysIysArgAlaIleLysAlaSerSerSerAspAspGluGly	689
Db	6361	GACCAGCTGCGCGCAGGCGGAGCGGCCCGCGCGCGCTGAGGAGCGGAGAGGCC	6420
Qy	690	-----GlyProArgThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLys	706
Db	6421	CGGCTGACGCGGAGGCTGAGGCGCGCAGGCCCGCGCAGCTGGAAGAGCGCGAGCGG	6480
Qy	707	AspIysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGln	726
Db	6481	CTGAAGCAGTCG-----CGAGGAGGACGAGCAGCGCCCGCGCTCAG	6522
Qy	727	GlySerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrp	746
Db	6523	GCACAGCGGCTCGACAG-----	6540
Qy	747	GluSerPheLysArgLeuValThrProArgLysLysSerLysLeuGluGluLys	766
Db	6541	-----AAGTCGCAAGGAGCGCGCAGCAGAGCGC	6570
Qy	767	AlaGluAspSerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSer	786
Db	6571	GCGCGCGCGCACAGCGCGGACGAGCGCGCTCGCGCAGAAAGCAGCAGCTGACGCGGAG	6630
Qy	787	TrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGln	806
Db	6631	ATGGAAGAAGCATAGAAATTCGCCGAGCAGACGCTCGCGCAGAGGCGCAGGTGGACAG	6690
Qy	807	GluGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnVal	826
Db	6691	GAGCTGACAACACTGCGG-----CTGCAGCTGGAGGACCGAC-----	6729
Qy	827	ProAlaValValProLeuSerGluTyrAsnAlaValGluArgGlu-----Lys	842
Db	6730	-----CACCAAGAAGAACCTGCTGGCAGCAGGAGCTGCACGCGCTCAAG	6771
Qy	843	MetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaVal	858
Db	6772	GCGAGGCCACGAGGCGCGCACCCACCGCAGGCTGGAGGTGGAGGAGCTCTTCGCGT	6831
Qy	859	TyrValSer---GluGluLeuSerLysThrLeuValHisThrValSerValAlaValIle	877
Db	6832	CGCGTCGAGATGGAGGAGCTGACCAAG-----	6858
Qy	878	AspGlyThrArgAlaValThrSerValGluLysArgSerProSerTrpIleSerAlaSer	897
Db	6859	-----CTCAAGGCACGCATCGCAGGCTGAGAACC CGCAGCTCATCTTCGCTG	6912
Qy	898	ValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProVal---GluGluVal	916
Db	6913	ANTACGCAAGCGCTTCCTGCAGGAGGAGGCTGAGAGATGAACAGGTGGCGGAGGAGGCC	6972
Qy	917	ThrGluLysAspIleLeAlaGluGluThrProValLeuThrGln-----Thr	932
Db	6973	GCGCGGCTGAGTGGCGGCCCAAGAGGCTGCGCAGCTGCGCAGCTGGCAGAGGAGAC	7032
Qy	933	LeuProGluGlyLysAspAlaHisAspAspMetValThrSerGluValAspPheThrSer	952

Db	7033	CTGGCACAGCAGCGCGGCTTGGCAGAGAGATGCTCAAGGAGAGATGTCAGGCGGTGCAG	7092
Qy	953	GluAlaValThrAlaThrSerGluAlaLeuArgThrGluGluValThrGluAla	972
Db	7093	GAGGCCACGCACTCAAGGCTGAGCGGAACCTGCTGCAGCAGCAGCAAGGAGCTTGGCGCAG	7152
Qy	973	SerGlyAlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerPro	992
Db	7153	GAGCAGCGCGCGCTGTCAGGAG	7176
Qy	993	AspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGlu	1012
Db	7177	GACAGGAGCAGATGCGCAGCAGCTGGCGGAGGAGACGCGAGGCTTCCACGCGACGCTG	7236
Qy	1013	GluGluGluArgGlnThrGlnAlaLeuGlnAlaValAlaAspLysValLys	1030
Db	7237	GAGCCGAGCGCAGCGCAGCTGGAGATGAGCGCTGAGGCTGAGCGCTCAAGCTCGT	7296
Qy	1031	-----GluGluSerGlnValProAlaThrGlnThrValGlnArgThrGlySer	1046
Db	7297	GTGCGCCAGATGATGCGCAGCCAGCGCCGCTGAGGAGGACGCCACGCGCTTCCGGAAG	7356
Qy	1047	LysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerGluLys	1066
Db	7357	CAGCGGAGAGATCGGTGAGAAGCTGCACCCACCGAG-----CTGCCACCCAGGAG	7410
Qy	1067	GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGln	1086
Db	7411	AAGGTGACCTGTGCAGACACTGGNGATCCAGCGACAGCAGAGTGACCATGTCGCGAG	7470
Qy	1087	GlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrAlaAsp	1106
Db	7471	-----CGCCTCGCGAGAGCCATCGCTGAG	7494
Qy	1107	ValAspHis-----ValAlaThrCysGlnValLleLysLeuGlnGlnLeu---MetGlu	1123
Db	7495	CTGAGCGGTGAGAGGAGAGAGTCCAACAGGAGGCCAACTGCTGCAGCTCAAGTCTGAG	7554
Qy	1124	GlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThr	1143
Db	7555	GAGATGCAGCGGTGCAGCGGAGCAGCTGCTGCAGGAGAGCGAGGCCCTCGCAGCAAGC	7614
Qy	1144	ProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGln	1163
Db	7615	TTCTCTCTGAAAAGGACACGCTG-----CTACAGCGGAGCGCTTCATCGACGAGGAG	7668
Qy	1164	AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThr	1183
Db	7669	AAGCCCAAGCTG-----GAGCAGCTCTTCCAGGACGAGGTGCCCAAG	7710
Qy	1184	AlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGly	1203
Db	7711	GCACAGCAG-----CTGCGTGAGGAG-----	7731
Qy	1204	GluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaVal	1223
Db	7732	CAGCAGCGCAGCAGCAGCAGATGGAGCAGGAACGCGCAGCGCTGGTGGCAGCATGGAG	7791
Qy	1224	ProValLeuAlaLysThrGluValGlyGlnGluGly-----GluVal	1237
Db	7792	GAGCGCGCGCGCGCAGCATGAGCGCGGAGGAGCGTGGCGCGCAAGCAGGAGGAGCTG	7851
Qy	1238	AspTrpLeuAspGlyGluLysValLysGluGluGlnGluValPheValHisSerGlyPro	1257
Db	7852	CAGCAGCTGGAGCAGCAGCGG-----CGCGACGAGGAGAGCTGCTGGCTGAGGAGAACCCAG	7908
Qy	1258	AsnSerGlnLysAlaAlaAspValThrTyAspSerGluValMetGlyValAlaGlyCys	1277
Db	7909	AGGTGCTGTGACGAGCTGTCAGCTCTCGGAGGAGCAGCACCGCGCGCGCTGCGCATCA	7968
Qy	1278	GlnGluLysGluSerThrGluValGlnSerLeuSer-----	1289
Db	7969	GAGGAGGTCACTGCTCGCAGGTGGTGTGCGCAAAAGACCTGCCCCAATGGCGGGGATGCA	8028

Db 3261 ATACTCAGAAAAATGTGTGACTTCTTAAGTGTCCCAAGAAATAGAAATTTATTTTAAACAGAC 3320
 Qy 363 nValGlyAspLeuGluAla -----SerSerGluGluLysCysAl 376
 Db 3321 AATATCTTGAGAAAGAGTGGAACTACAGCATATATAAGGAAGATTTGGAAGAAAAG----- 3375
 Qy 376 aProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValValAlaG1 396
 Db 3376 -----TGCGCAGCTGAA-----GAGCAATTCAGGCTCTGTGTCAACA 3413
 Qy 396 uValHisValSerThrValGluLysThrGluGluGlnGluGlyGlyGlyGluAlaG1 416
 Db 3414 GATGAATCAGACCTTCGACAGATAAACAAACCAA----- 3447
 Qy 416 uGlyGlyValValGluGlyThrGlyGlySerLeuProGluLysLeuAlaGluPr 436
 Db 3448 -----ATAGATTTTGCTCCAAGCAGCAAAATCAGTGAAAA 3479
 Qy 436 oGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCy 456
 Db 3480 CCAAGCAATTTATCCAGAAG-----TTAATCACAAGTAGTAACACGGA 3518
 Qy 456 sValSerGlyGlyAspHisThrGlnLeuThrAsp-----LeuSerProGluG1 472
 Db 3519 TGCAGGTGATGGGACTCCGTAGCACTTGTAAAGGAAACAGTGTGTATAGTCCACCTTG 3578
 Qy 472 uLysThrLeuProLysHisProGluGlyLeuValSerGluValGluMetLeuSerSerG1 492
 Db 3579 TACAGGTAGTAGTGAACACTGGAAACCAGAACTAGAGAAGATATACTGSCCCTTGAAAA 3638
 Qy 492 nGluArgG1LeLysValGlnGlySerProLeuLysLysLysLeuPheSerSerSerGlyLeuLy 512
 Db 3639 AGAAAAGGAGCAACTTCAAAGAAGCTTACAGGAAGCCTTAACCTCCCGCAAGGCAATCT 3698
 Qy 512 sLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyGlyGluProG1 532
 Db 3699 TAAAAAGGCACAGAGAAAGAAAGACATCTCAG-----GAGGAGCTAAA 3743
 Qy 532 yGluTyGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSe 552
 Db 3744 GCAACAGAAAGATGACTATATCGCTTGCAAGAACAGTTTGTATGAGCAAGCAAGGAAAA 3803
 Qy 552 r-----SerAl 554
 Db 3804 TGAGAAATATTGGAGACCAGCTAAGGCAACTCCAGATTCAAGTAAGGGAATCCATAGACGG 3863
 Qy 554 aSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPr 574
 Db 3864 AAACTCCCAGACAGCAGCAGGAGGATCGTGTCTTCCACCTCCAGGTTTAGAAGAAC 3923
 Qy 574 oGlnAspGlyGluAlaGluGluGlyThrThr----- 594
 Db 3924 TTTATTCAAAGCCACAGAACAGATCACACTCAACCTGTTTTAGAGTCCAACCTGTGCC 3983
 Qy 585 -----SerAspGlyGluLysLysArgGluGlyIleThrProTrpAlase 599
 Db 3984 AGACTGCGCTTCTCATTTCTGAAGATGCGAGTGTCTGCGAGGCGGAACCTCTCTGCCCA 4043
 Qy 599 rPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysG1 619
 Db 4044 GATTAAAGCCACAGCTG-----AAGGA 4064
 Qy 619 uGluGluLeuGluLysValLys---SerAlaThrLeuSerSerThrAspSerThrValSe 638
 Db 4065 AATAGAGCTGAGAAAGTAGAGTTAGAAATTTAGAAAGTTAGTTTACAACAAGTAGCTTAC 4124
 Qy 638 rGluMetGlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysAr 658
 Db 4125 TAAAAAATCAGAAAGAGGTATTTTCAGTTTACAAAGCAGATATAATAAACACAGGGTTTAGAAAT 4184
 Qy 658 gArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysAr 678
 Db 4185 CGAGAGTCTTAAGACAGATATCCCAAGCTGAAGTCCATGCCGAAGAGCCTGCAGCAGAA 4244

Qy 678 gAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSe 698
 Db 4245 ATTT-----CAAAAGCAGCCAACTACAAATTTGCTGGCCTAGA 4280
 Qy 698 rHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlase 718
 Db 4281 ACATCTAAGAGAATTCGAACCTTAACCTGGATGAACCTG-----CAAAAACCTCAATAG 4331
 Qy 718 rThrGlnGluGlnAspGlnAlaGlnGlySer-----SerSerProGluLuproAlaG1 735
 Db 4332 CAAAAAGGAAGAGAGCTTAGCTACCTTTCTGGACAACTTAGTGAGAAAGACAGCTCT 4391
 Qy 735 ySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeuValThrPr 755
 Db 4392 CACTAAATACAGACAGAGATAATAGAACAAAGAGATTTAATTAAGGCTCTGCATACA-- 4449
 Qy 755 oArgLysLysSerLysSerLysLeuLysAlaGluAspSerSerValGluGlnLe 775
 Db 4450 -----CAGCTAGAAATGCAAGCAAGAGCATGATGAGAGATTAAGCAGCT 4496
 Qy 775 uSerThrGluIle-----GluProSerArgGluGluSerTrpVa 788
 Db 4497 ACAGGTGGAACTTTGTGAATGAACCAAAACCCAGAGAGATTGGAGAGAAAGT----- 4551
 Qy 788 lSerIleLysLysPheIleProGlyArgLysLysArgAlaAspGlyLysGlnGlu 808
 Db 4552 -----AGACAAAGCAACAAATACAAAGGAAACTGCAAGC 4586
 Qy 808 nAlaThrValGluAspSerGlyProValGluLeuAsnGluAspAspProAsnValProAl 828
 Db 4587 TGCCCTTATTTCCCGAAAA-----GAAGCACTTAAAGAAACAAAGTCTCCAAGA 4637
 Qy 828 aValValProLeuSerGluTyAsnAlaValGluArg----- 840
 Db 4638 GGAATTTCTTTGGCCAGA---GGTACCATTGAACGCTCTCACCAGTCTCTGCGCAGATGT 4694
 Qy 841 -GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuLysAlaValTyVa 860
 Db 4695 GGAAGCCCAAGTTTCTGCTCAAAATAAAGAAAGATACGGTCTTTAGGAAGGTTAGCTCT 4754
 Qy 860 lSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyTh 880
 Db 4755 TCTTCAAGGAAGAAAGACAAACTCTATTACAGAAATGGACAGTCTTTTATGGAATCA 4814
 Qy 880 rArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrG1 900
 Db 4815 GAGTCTCAGCAGCTCTCTGTGAA-----AGTCTAAAACTAGCTCTAGAGGGTCTTACTGA 4868
 Qy 900 uProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAs 920
 Db 4869 AGACAAAGGAAAAGTTAGTGAAGGAAAT-----GAATCTTTGAAATCTTC 4913
 Qy 920 pIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHi 940
 Db 4914 TAAGATTGCAAGAAAGTACT-----GAGTGGCAAGAGAAACA 4949
 Qy 940 sAspAsp-----MetValThrSerGluValAspPheThrSerG1 953
 Db 4950 CAAGGAGCTACAAAAGAGATGAAATTTCTCTGCGAGTCTTATGAGATGTTAGTAATGA 5009
 Qy 953 uAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGlu----- 968
 Db 5010 AGCAGAAAGGATTCAGCATGTGTGGAGCTGTGAGGCAAGAGAAACAAGAACTGTATGG 5069
 Qy 969 -----ValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaVa 985
 Db 5070 CAAGTTAAGACACAGAGGCAACAAGAAAGGAGACAGAAAAAGCAGCTGCAGGAAGCTGA 5129
 Qy 985 lSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG1 1005
 Db 5130 GCAAGAAATGGAGGAATGAAAAGAAAGATGAGAAAGTTT-----GCTAAATCTTAAACA 5183

QY 59 rSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGlnGluGlnGluGln 79
 DB 2493 CCATGCAACCTCTCTCACTGACAG-----ATCCATAGTCTCAGCATAGAACCCAAATC 2546
 QY 79 yGlnGluGluGluValValAlaGluGlnGlyAspValGlnGlnGlnGlnGlnGlnGln 99
 DB 2547 TAAAGATGTGAAATTT-----GAAATTTTACAGATGAATGATGTGAG-- 2595
 QY 99 uLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValAlaGluAspIleThrLysAs 119
 DB 2596 -----CTTCAGTCTTCAGCAGAGATACCTGTATGAAGGCTGCAAGGCCA 2642
 QY 119 pGlyGlnGluGluThrSerGluIlelele-----GluGlnIleProLaserGluAs 136
 DB 2643 GCTGCAAAATTAAGGAAGTGAAGTCTTGAAGGGGAGAGACGTGTGAAGGCATATCTCAAG 2702
 QY 136 nAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLys 156
 DB 2703 TAAAGTGAAGAACTGTCCAGGCTCTTTCACAGAGGAAGAACTTGAATA----- 2751
 QY 156 sValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAs 176
 DB 2752 -----ACAAAATGAT----- 2763
 QY 176 pThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValG 196
 DB 2764 -----CAGCTCTTACTAGAGAAAAGAGAGATGTGGAACCTCCACAAAACCATCGA 2816
 QY 196 yAlaGlyAspHisGlnGluProSerValGluThrAlaValAlaGlyGluSerAlaSerLysG 216
 DB 2817 GGAGAGGATCAACAGTGCAGAAATCAGCTTTAGTAGTACGTGAGAAAATGCTTCAGCT 2876
 QY 216 uSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGlnGlnSerSe 236
 DB 2877 TAATGAAGAGAAGTTTCTCTCGGGTTGAATTAAGACTCTTAAGAA----- 2925
 QY 236 rThrGluIleProLeuGlnAlaGluSerAspGlnAlaGluGluGluAlaLysAspG 256
 DB 2926 -----CAGCTAAATTTATTTATCCAGAGCTGAGGAAGCAAAAGAGAGCAGCGTGAAGAAGA 2981
 QY 256 uGlyGluGluLysGlnGluLysGluProThrLysSerProGluSerProSerProVa 276
 DB 2982 TAATGAAGTTTCTTCTGGCTTAAACAAAATTAAGATAGATGAGCCAGCAGGACAAAT 3041
 QY 276 lAsnSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpAr 296
 DB 3042 AAGTAGGAAGAA-----CTTCACATGAATTTGACCTTCTGAA 3080
 QY 296 glyLysThrSerPheLysLysSerLys-----GluAspAs 308
 DB 3081 GAAAGAAAATGAGCAGAGAAAGAGAAAGCTCCAGGAGCTCTTTATTAACAGAAAGGAGCT 3140
 QY 308 pLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluGluGly 328
 DB 3141 TCTGCAAGAGTCAAGTAGTGAAGAGAAATTAAGCAACTTGAAGAGATGAATCTAAGAA 3200
 QY 328 sGluLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaAr 348
 DB 3201 AGAAATCCCACTCAGTGAGACTGAGAGGGGAGAGAGTGAAGAGAGATGAAGAAACAAAGA 3260
 QY 348 gLeuSer-----AlaAspTyrGluLysValGluLeuProLeuGluAspG 363
 DB 3261 ATACTCAGAAAAATGTGTGACTTCTAAGTCCCAAGAAATAGAAATTTATTTAAACAGAC 3320
 QY 363 nValGlyAspLeuGluAla-----SerSerGluLysLysCysAl 376
 DB 3321 AATATCTGAGAAAGAGTGGAACTCAGCATATATAAGGAGGATTTGGAAGAAAG----- 3375
 QY 376 aProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValValAlaG 396
 DB 3376 -----CTGGCAGCTGAA-----GAGCAATTCAGGCTCTGTGTCAAACA 3413

QY 396 uValHisValSerThrValGluLysThrGluGluGlnGlnGlyGlyGlyGlyGluAlaG 416
 DB 3414 GATGAATCAGACTTTCGAAGATAAAACAAACCA----- 3447
 QY 416 uGlyGlyValValValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluPr 436
 DB 3448 -----ATAGATTGCTTCCAGAGCAGAAATCAGTGAAAA 3479
 QY 436 oGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCy 456
 DB 3480 CCAAGCAATATCCAGAAG-----TTAATCACAAGTAACACGGA 3518
 QY 456 sValSerGlyGlyAspHisThrGlnLeuThrAsp-----LeuSerProGluG 472
 DB 3519 TCAGATGTATGGGACTTCGTAGACATCTGTAAGGAAACAGTGGTATAGTCCACCTTG 3578
 QY 472 uLysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerG 492
 DB 3579 TACAGGTAGTAGTAACACTGGAACACCAGAACTAGAGAAAAGATACTGGCCCTTGAANA 3638
 QY 492 nGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuL 512
 DB 3639 AGAAAAGGAGCAACTTCAAAAGAACGCTACAGGAAGCCTTAACCTCCCGCAAGCAATTCT 3698
 QY 512 sLysLeuSerGlyLysLysGlnLysGlyArgGlyGlyGlyGlyGlyGlyGluProG 532
 DB 3699 TAAAAAGGCACAGGAGAAAGAAACATCTCAG-----GAGGAGTAAA 3743
 QY 532 yGluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSe 552
 DB 3744 GCAACAGAAAGATGACTATATTCGTTGCAAGACAGATTTGATGAGCAAGCAAGGAAAA 3803
 QY 552 r-----SerAl 554
 DB 3804 TGAGATATTGGAGACCAGCTAAGCACTCCAGATTCAAGTAAGGAATCCATACGCG 3863
 QY 554 aSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPr 574
 DB 3864 AAAACTCCCAAGCAGACAGCAGCAGGAATCGTGTCTTCCACTCCAGGTTTAGAAGAACC 3923
 QY 574 oGlnAspGlyGluAlaGluGluGlyThr----- 584
 DB 3924 TTTATTCAAGCCACAGAACAGCATCACACTCACTGTTTGTAGCTCAACTGTGTCGCC 3983
 QY 585 -----SerAspGlyGluLysArgGlyIleThrProTrpAlaSe 599
 DB 3984 AGACTGGCCTTCTCATTCTGAAGATCCGAGTCTCTGCGAGCGCGGAACCTCTCTGCCCCA 4043
 QY 599 rPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysG 619
 DB 4044 GATTAAGGCCAGCTG-----AAGGA 4064
 QY 619 uGluGluLeuGluLysValLys-----SerAlaThrLeuSerSerThrAspSerThrValSe 638
 DB 4065 ATAGAGGCTGAGAAGTAGAGTTAGAAATTGAAGTTAGTTCTTACCAACAGTCAAGCTTAC 4124
 QY 638 rGluMetGlnAspGluValLysThrValGlyGluGlnLysProGluGluProLysAr 658
 DB 4125 TAAAAATCAGAAAGAGTATTTTCAGTTACAGAGCAGATATAATAACACAGGGTTTAGAAAT 4184
 QY 658 gArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysAr 678
 DB 4185 CGAGAGTCTAAAGACAGTATCCCATGAAGTCCATGAGTCCGAAAGGCTTCGAGCAGAA 4244
 QY 678 gAlaArgLysAlaSerSerSerAspAspGluGlyProArgThrLeuGlyGlyAspSe 698
 DB 4245 ATTG-----GAAAGCAGCCCACTACAAATTTGCTGGCCTAGA 4280
 QY 698 rHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProLase 718
 DB 4281 ACATCTAAGAGAAATGCAACCTTAACCTGATGAACTG-----CAAAAACCTCATAG 4331
 QY 718 rThrGlnGluGlnAspGlnAlaGlnGlySer-----SerSerProGluProAlaG 735

```

Db 4332 CAAAAGGAAGAGCTTACCTCTTGTGACAACTTAGGAGAAAGAACCTCT 4391
QY 735 ySerProserGluGluGluGluValSerThrTrpGluSerPheLysArgLeuValThrPr 755
Db 4392 CACTAAATATACAGACAGATTAATAGAACAGAAAGATTATTAATGAAGCTCTGATACA-- 4449
QY 755 oArgLysSerLysSerLysLeuGluGluValSalAGluAspSerSerValGluGluLe 775
Db 4450 -----CAGCTAGAAATGCAAGCAAGCAAGCATGATGAGAGGATTAAGACGCT 4496
QY 775 userThrGluLe-----GluProSerArgGluGluSerTrpVal 788
Db 4497 ACAGGTGAACTTTGTGAATGAAGCAAAACCAAGAGATGGAGAGAAAGT----- 4551
QY 788 lSerLysLysPheLeProGluYArgArgLysValArgLysAspLysGlnGluGlu 808
Db 4552 -----AGAGCAAGCAAAATGCAAAAGGAACTGCAAGC 4586
QY 808 nAlaThrValGluAspSerGlyProValGluLeaGluAspAspProAsnValProAl 828
Db 4587 TGCCCTTATTTCCCGAAAA-----GAGCACTAAAGAAACAAAGATCTCCAAAG 4637
QY 828 aValValProLeuSerGluTyrAsnAlaValArg----- 840
Db 4638 GGAATGTCTTTGGCCAGA---GCTACCATGAACTCTTCACCAAGTCTTCGCAATGT 4694
QY 841 -GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVa 860
Db 4695 GGAAGCCCAAGTTTCTGCTCAAAATGAAGAAAGATACGGTCTTGAAGAGTTGACTCT 4754
QY 860 lSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValLysGlnGly 880
Db 4755 TCTTCAAGAGAAAGAAAGACAAACTCATTCAGAAATGACAGAGCTTTATGAAATCA 4814
QY 880 rAlaGluAlaValThrSerValGluGluLysArgSerProSerTrpLysSerAlaSerValThrG 900
Db 4815 GAGCTTCAGACAGCTCTCTGTA---AGTCTAAACTAGCTCTGAGAGGCTTTACTGA 4868
QY 900 uProLeuGluHisThrAlaGlyLysLysLysMetProProValGluGluValThrGluLysAs 920
Db 4869 AGCAAGAGAAAGATTAGTGAAGAAATT-----GAATCTTTGAAATCTTC 4913
QY 920 pIleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
Db 4914 TAAGATTGCGAAAGTACT-----GAGTGGCAAGACAGAAAC 4949
QY 940 sAspAsp-----MetValThrSerGluValAlaAspPheThrSerG 953
Db 4950 CAAGGAGCTACAAAAGATGATGAATTTCTTGCAGCTCTTATGAGAAATGTAGTATGA 5009
QY 953 uAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGlu----- 968
Db 5010 AGCAGAAAGGATTCAGCATGTGTGAGAGCTGTGAGGCAAGAGAAACAAGAACTGATGG 5069
QY 969 -----ValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaVal 985
Db 5070 CAAGTTAAGAAAGCAGAGAGCAAAAGAAAGAGACAGAAAGACAGTTGCGAGAACTGA 5129
QY 985 lSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005
Db 5130 GCAGAAATGAGGAATGAAAGAAAGATGAGAAAGTTT-----GCTAAATCTTAAACA 5183
QY 1005 uSerGlyValLeuAspThrThrGluGluGluGluValArgGlnThrGlnAlaLysLeuGlnAlaVal 1025
Db 5184 GCAGAAATCTCTAGACTGAGAGAGAAAGATGACCGCTTAAAGGAGAGGTGACCTGC 5243
QY 1025 lAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
Db 5244 AGGAGATTACAGCTAAAGAG-----TGATGAAACACTTCTTTCTTCCAA 5288
QY 1045 ySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065

```

```

Db 5289 TGCAGCATGAAGAGAACTTGAAGGGTCAAAATGAGATATGAACCTTTTAAAGA 5348
QY 1065 u-----LysGluLysAspValMetProLysGlyProValGlnGluAlaG 1080
Db 5349 GTTTCAGCTTTAATGTCTGGAAGAAAGACTCTTAAGTAA---GAGGTTCAGAT----- 5400
QY 1080 yAlaGlnLysLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLysGluVal 1100
Db 5401 -TTAAAGCATCGATAGAGAGATATGTATCTTAACAGACTTAACCTAGAGGCCACCGAGAA 5459
QY 1100 lProGluValThrAlaAspValAlaAspHisValAlaThrCysGlnVal-----Lys 1117
Db 5460 ACATGATTAACCAAGAAATGTCATGAAAGAGGAACACAGCTTATACAGAGTGAAGTGA 5519
QY 1117 lLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSe 1137
Db 5520 AGAGCAAGCTCTCTGAGATATGAGCAAGAACTTCAGTTTCAGAAATGCGTTCCATCAGC 5579
QY 1137 rGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnAla 1157
Db 5580 GAAGAGT-----GCCAACCTGCTGTAGTAAAGATTTTCACTCAGATGAATTTAA 5633
QY 1157 p-----GluThrLysSerGlnAspSerLysAlaThrAlaAlaValArgLys 1174
Db 5634 TAACATACCTACAGCAGATTTGATCAGCTCAAGAAAGAAATGCTGATTAAGAGAGAGA 5693
QY 1174 rGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAs 1194
Db 5694 GCAGAAAGCAAGAAATTTAGCCAGCACTTTAGAAATGAGAAA---AATACCTTACTAG 5750
QY 1194 nAsnValProAlaGlnGluGluHisGlyGluGluProGluYArgAspValLeuGluProTh 1214
Db 5751 TACGATATCAACAAAGAT-----GATGA-----CTAAATATGCT 5786
QY 1214 rGlnGlnGluLeuThr-----AlaAlaAlaValProVal 1225
Db 5787 TCAGAGGAGAAATGAACAAATGAACCTGTTAAATCAGCAAAATCCAAAGAACTCCAG 5846
QY 1225 lLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysVa 1245
Db 5847 ACTTACCAAACTAAAGAGAGACAGCAAGAGAAAGAAATGATTTGGAAAGAGAGCTTAT 5906
QY 1245 lLysGluGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspVa 1265
Db 5907 GAATCAATTAGCAAGACTT-----AATGAAGCACTTGGAAATTTCTGTCAGAGATG 5957
QY 1265 lThr-----TyrAspSerGluValMetGlyValAlaGlyCy 1277
Db 5958 TACAGATGCCAAATTAATAAATAGAGCTATTGGAATCTGAATGAAGAACTTAATAAGTG 6017
QY 1277 rGlnGluLysGluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluTh 1297
Db 6018 TGTGAGTGAATTGGAAGAAAGAAAGACAGCACTTACTCAAGAAAGAACTAAAGTGAATC 6077
QY 1297 rAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGluGluGluG 1317
Db 6078 ACAAATTAACAAAGGAAATTTGGAGAAATACAGAGTCTCAGAAAGAACCCGGAATTA 6137
QY 1317 nGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMe 1337
Db 6138 A-----AGCCATGCAAAAGAACTTCAGAAATCTTAA 6170
QY 1337 rProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyLysSerProSerLeuProAs 1357
Db 6171 AGAAAAACAAAGAAATGAAGACACTA-----CAGAAGCACTGCATCAGTA 6218
QY 1357 rGlnAspLysAlaGlyCysLysGluValGlnValGlnSerLeuAspThrThrValThrG 1377
Db 6219 TCAGAGAAATTAAGTGTCTTGGAGAGAACTGTTAAGCTTGAAGATTTGTTCAAACTGA 6278
QY 1377 nThrAlaGluAlaValGluLysValLysGluThrValLysSerGluThrGlyGlyLys 1397
Db 6279 ATCTCAAAAGATTTGAAATTAACCAAGAAAT-----CTGGCTCAAGAGTTGAACA 6332

```

QY 1397 rProGluCysValGlyAlaHis-----LeuLeuProAlaGluLysSe 1411
 Db 6333 CCACAAGAGGCAACAGCAGAAATTAGTAGCTTCAAAGTCTCTGTAGATGACACTCAAAG 6392
 QY 1411 rSerAlaThrGlyGlyHisrPrThrLeuGlnHisAlaGluAspThrValProLeuGlyPr 1431
 Db 6393 TGAAGCAGCA-----AGGTCCTAGCAGACAACTCAAGTTGAAAAA 6434
 QY 1431 oGluSerGlnAla-----GluSerIleProIleValThrProAlaProGluSerTh 1449
 Db 6435 GGAACCTTCAGTCAATAAAGAAATCAGTTAAAGCCAGCAACAAAGGATGAA----- 6489
 QY 1449 rLeuHisProAspLeuGlnGlyGluLysSerAlaSerGlnArgGluArgSerGluGlu 1469
 Db 6490 -----GATCTTGAGCGAAGACTGGAAACAGCGAGAGAGAACCCCTGAAAGAGAA 6539
 QY 1469 uAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysVa 1489
 Db 6540 GAAGAATATGCAAGAGAACTGGATGCTTGGCGAGAGAAAAAGTCCACTTGGAGAGAC 6599
 QY 1489 lLeu-----LysAlaGluProGluIleLeuGluLeuGluSe 1501
 Db 6600 AATTGGAGAGATTTCAGTTACTTTTGAACAAGAAAGCAAGCAAGTTCAGCACTTCAGGA 6659
 QY 1501 rLysSerAsnLysIleVal 1507
 Db 6660 AACTTGGACAGTACTGTG 6678

RESULT 14
 US-09-724-676-7240
 ; Sequence 7240, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 7240
 ; LENGTH: 8189
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676-7240

Alignment Scores:
 Pred. No.: 1 98e-09 Length: 8189
 Score: 415.00 Matches: 316
 Percent Similarity: 38.17% Conservative: 305
 Best Local Similarity: 19.42% Mismatches: 669
 Query Match: 5.14% Indels: 337
 DB: 5 Gaps: 60

US-09-902-432-4 (1-1596) x US-09-724-676-7240 (1-8189)

QY 26 LeuSerGly-----HisGlyProAlaAlaGluAlaSerGlyAlaAla----- 39
 Db 2404 CTCTCAGGTGAAGAACTTAGCATGGTAACAGAAATTGAGGGCTCAGTAAAGCACTGGA 2463
 QY 40 GlyAspPro-AlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSe 59
 Db 2464 AATGAACCTTGCAGAGCAGAGAGGCAAGCAAGCACTTGATTATGAAAGC---CAAACTGC 2520
 QY 59 rSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGlnGlnGluGlu 79
 Db 2521 CCATGCAACCTCTGCTACTGAACAG-----ATCCATAGTCTCAGCATAGAAGCCAAATC 2574
 QY 79 yGlnGluGluGluValValAspGluAspGluValGlnArgGluSerGluAspValArgG 99
 Db 2575 TAAAGATGTGAAATTT-----GAAGTTTACAGATGAATGATGTGTCGAG-- 2623
 QY 99 uLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAs 119

Db 2624 -----CTTCAGTTTTCTGAGCAGAGTAGTACCTCGATAAGAAAGCCTCCAAAGCCA 2670
 QY 119 pGlyGlnGluGluThrSerGluIleIle-----GluGlnIleProAlaSerGluAs 136
 Db 2671 GCTGCAAAATAAGAAAGTGAAGTCTTCAGGGGCGCAGAACCTGTGAAGCATATCTCAAG 2730
 QY 136 nAenValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLy 156
 Db 2731 TAAAGTGAAGAACTGCTCCAGGCTCTTTCACAGAAGAACTTGAATA----- 2779
 QY 156 sValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAs 176
 Db 2780 -----ACAAAATGGAT----- 2791
 QY 176 pThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValG 196
 Db 2792 -----CAGCTTCTACTAGAGAAAGAGAGATGTGGAACCTCCACAAACCATCGA 2844
 QY 196 yAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysG 216
 Db 2845 GGAGAAGGATCAACAAGTGCACAGAAATCAGCTTTAGTATGACTGAGAAAATGGTTCAGCT 2904
 QY 216 uSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSe 236
 Db 2905 TAATGAAGAGAAAGTCTTCTTGGGTTGAAATTAAGACTCTTAAAGAA----- 2953
 QY 236 rThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspG 256
 Db 2954 -----CAGCTTAAATTTATATCCAGAGCTGAGGAAGCAAAAAGACGCGTGAAGAGA 3009
 QY 256 uGlyGluGluLysGlnGluGlyProThrLysSerProGluSerProGluSerProSerProVa 276
 Db 3010 TAATGAAGTTCTTCTGGGCTTTAAACAAATTTATGATGAGATGAGCCAGCAGGACAAAT 3069
 QY 276 lAenSerGluThrThrSerSerPheLysLysPheThrHisGlyTrpAlaGlyTrpAr 296
 Db 3070 AAGTAAGGAAGAA-----CTTCAGCATGAATTTGACCTTCTCGAA 3108
 QY 296 gLysLysThrSerPheLysLysSerLys-----GluAspAs 308
 Db 3109 GAAAGAAATAGCAGAGAGAAAGAAAGCTCCAGGCGCTCTTTTAAACAGAAAGGAGCT 3168
 QY 308 pLeuGluThrAlaGluLysArgLysGlnGluAlaGluLysValAspGluGluGlu 328
 Db 3169 TCTGCAAGAGTCAGTAGATTGGAAGAGAAATTAGCCAACTTGAAGATGAATCTAAGAA 3228
 QY 328 sGluLysThrGluProAlaSerGluGlnGluProAlaGluAspThrAspGlnAlaAr 348
 Db 3229 AGAAATCCCACTCAGTGAGACTCAGAGGGGAGAGAGTGAAGAGATAAAGAAACAAAGA 3288
 QY 348 gLeuSer-----AlaAspTyrGluLysValGluLeuProLeuGluAspG 363
 Db 3289 ATACTCAGAAATATGTGCTACTTCTAAGTGCAGAAATAGAAATTTATTTAAACAGAC 3348
 QY 363 nValGlyAspLeuGluAla-----SerSerGluGluLysCysAl 376
 Db 3349 AATACTGAGAAGAGAGTGAAGTACAGCATATAGGAAGGATTTGGAAGAAAG----- 3403
 QY 376 aProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValValaG 396
 Db 3404 -----CTGCGAGCTGAA-----GAGCAATTCAGGCTCTGCTCAACA 3441
 QY 396 uValHisValSerThrValGluLysThrGluGluGlnGlyGlyGlyGlyGluAlaG 416
 Db 3442 GATGAATCAGACTTGCAGAGATAAACCAACCA----- 3475
 QY 416 uGlyGlyValValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluPr 436
 Db 3476 -----ATAGATTGCTCCAGCGAGAAATCAGTGAAAA 3507
 QY 436 oGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCy 456

```

Db 3508 CCAAGCAATTATCCAGAG-----TTATCAAGATTAACGGA 3546
Qy 456 eValserGlyIAspHisThrGlnLeuThrAsp-----LeuSerProGluG 472
Db 3547 TCCAAAGTGAAGGAGCTCCGTAGCACTGTAAAGAAAGAGTGGTATGATCAAGCTTG 3606
Qy 472 uLysThrLeuProLysHisProGluGlnIleValSerGluValGluMetLeuSerG 492
Db 3607 TACAGGTAGTACTGAACCTGGAAACCAAGACTAGAAAGAAAGATATCTGGCCCTTAA 3666
Qy 492 ngLysArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerGlyLeu 512
Db 3667 AGAAAAGAGCAACTTCAAAAAGAGCTACAGAAAGCTTAACCTCCGCAAGCAATCTCT 3726
Qy 512 sLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysArgGluProG 532
Db 3727 TAAAAAGCAAGAGAAAGAAAGAAAGACATCTCAAG-----GAGAGCTTAA 3771
Qy 532 yGluTyrgLnisIleHisThrGlnSerProGluSerAlaAspGluGlnLysGlyL 552
Db 3772 GCACAGAAAGATGACTATATCTTGCAAGAACAGTTGATGACCAAGCAAGGAAA 3831
Qy 552 r-----SerAl 554
Db 3832 TGAAGATATTGGAGACCACTAAGCACTCCAGATTCAAGTAAGGAATCCATAGCG 3891
Qy 554 aSerSerProGluGlnProGluGlnThrCysLeuGlnLysGlyProLeuGlnAlaP 574
Db 3892 AAAATCCCAAGACAGACAGCAAGAAATCGTGTCTTCCATCCAGCTTAAAGAAAC 3951
Qy 574 oGlnAspGlyGlnAlaGlnGlnGlyThr-----584
Db 3952 TTTATTCAAAAGCCACAGAACAGATCACTCACTGTTTAAAGTCCAACTTGCC 4011
Qy 585 -----SerAspGlyGlnLysLysArgGlnGlyLysThrProTriPase 599
Db 4012 AGACTGCGCTTCATTCATTCGAATGCAAGTGTCTGCAAGGCGAACTTCTTGCCA 4071
Qy 599 rPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysG 619
Db 4072 GATTAGGCCCAAGCTG-----AAGGA 4092
Qy 619 uGlnGluLeuGlnLysValLys-----SerAlaThrLeuSerThrAspSerThrVal 638
Db 4093 AATGAGGCTGAGAAAGTGAAGTGAATGAAGTTAAAGTTAGTTACAAAGAGCTTAC 4152
Qy 638 rGluMetGlnAspGlnValLysThrValGlnGlnGlnLysProGluGlnProLysAr 658
Db 4153 TAAAAATCAGAAAGAGTATTTCAGTTACAGAGCAGATTAATAACAGGTTTAGAAT 4212
Qy 658 gArgValAspThrSerValSerTriPaseLysLysValGlySerSerLysArg 678
Db 4213 CGAGACTTAAAGACAGTATCCATGAGTGAAGTCAATGCCGAAAGCTCGACAGAA 4272
Qy 678 gAlaArgLysLysSerSerSerAspAspGlnGlyGlyProArgThrLeuGlnLysP 698
Db 4273 ATTG-----GAAAGCAAGCACTCAAAATGCTGCTTGA 4308
Qy 698 rHisArgLysGlnGlnLysLysSerLysAspLysGlnGlnLysThrAspAlaValPro 718
Db 4309 ACATCTTAAGAGATTGCAACCTTAACCTGAGTGAAGT-----CAAAATCTTAAG 4359
Qy 718 rThrGlnGlnLysAspGlnLysLysSer-----SerSerProGluProAlaG 735
Db 4360 CAAAAGAGAAAGAGCTTACCTTCTGACAACTTGAAGAGAAAGACACACTCT 4419
Qy 735 ySerProSerGlnGlyGlnGlyLysLysThrTriPaseLysArgLeuValThrP 755
Db 4420 CACTAAATAATACAGACAGATTAATAGAAAGAAAGATTTAATTAAAGCTCTTGATCA 4477
Qy 755 oArgLysLysSerLysSerLysGlnGlnLysLysAlaAspSerSerValGlnLys 775
Db 4478 -----CAGCTAGAAATGCAAGCCAAAGACATGATGAGAGATTAAGACGCT 4524

```

```

Qy 775 userThrGluLe-----GluProSerArgGlnGlnGlnSerTrpVa 788
Db 4525 ACAGGTGAACCTTGTGAATGAAGCAAAAACCAAGAAAGATGGAGAAAGAGT----- 4579
Qy 788 lSerLysLysPheLeuProGlnArgArgLysLysArgAlaAspGlyLysGlnGln 808
Db 4580 -----AGAGCAAGCAACAAATTAACAGAAAGAACTCCAGC 4614
Qy 808 nAlaThrValGlnAspSerGlyProValGlnLysGlnLysAspProAlaValProAl 828
Db 4615 TGCCCTTATTTCCGAAAA-----GAAGCACTAAAGAAAGAAAGAAAGCTCCAGA 4665
Qy 828 aValValProLeuSerGlyTyrgAsnAlaValGlnArg-----840
Db 4666 GGAATGTCTTGGCCACA--GGTACCATTAAGAGTCTTCAACAGTCTGCGAGATGT 4722
Qy 841 -GlnLysMetGlnAlaGlnGlnLysThrGlnLeuProGlnLeuGlnLysAlaValTyrg 860
Db 4723 GGAAGCCCAAGTTTGTCTCAAAATTAAGAAAGATACGGTCTTAGAGAGTTAGCTCT 4782
Qy 860 lSerGlnGlnLeuSerLysThrLeuValHisThrValSerValAlaValLysArg 880
Db 4783 TCTTCAAGAAAGAAAGACAAACTCATTAACAGAAATGACAGGTCTTATGAAATCA 4842
Qy 880 rArgAlaValThrSerValGlnGlnLysArgSerProSerTriPaseAlaSerValThr 900
Db 4843 GAGTCTCAGCACTCTGTGA-----AGCTTAAACTAGCTTAAAGGCTTAAAGCA 4896
Qy 900 uProLeuGlnHisThrAlaGlnLysAlaMetProProValGlnGlnLysLys 920
Db 4897 ACACAAGAAAGAAAGTGAAGAAAT-----GAATCTTGAATCTTC 4941
Qy 920 rIleIleAlaGlnGlnLysProValLeuThrGlnThrLeuProGlnGlyLysAspAla 940
Db 4942 TAAAGTTGCAAGAAAGTACT-----GAGTGCAAGAGAAACA 4977
Qy 940 sAspAsp-----MetValThrSerGlnValAspPheThrSerG 953
Db 4978 CAAGAGCTACAAAGAAAGATGAATCTTCTGCAAGTCTTAAGAAATGTAATGA 5037
Qy 953 uAlaValThrAlaThrGlnThrSerGlnLysAlaThrArgThrGlnGln-----968
Db 5038 ACACAAGAAAGATTCAGCATGCTGTGAAAGCTGTGAGCAAGAAACAGAACTATAG 5097
Qy 969 -----ValThrGlnLysSerGlyAlaGlnLysThrAspMetValSerAla 985
Db 5098 CAAGTTAAGAAACACAGAGGCAACAGAGAGACAGAAAGCAAGTTGCAAGAGCTGA 5157
Qy 985 lSerGlnLeuThrAspSerProAspThrThrGlnGlnLysAlaThrProValGlnGln 1005
Db 5158 GCAGAAATTAAGAGAAATGAAGAAAGATGAGAAAGTTT-----GCTTAATCTTAAC 5211
Qy 1005 uSerGlyValLeuAspThrGlnGlnGlnLysArgGlnThrGlnAlaIleLeuGlnAla 1025
Db 5212 GCAGAAATCTTAGAGCTGGAAGAGAGATGACCGGCTTAGAGGAGGTCAACCTGCG 5271
Qy 1025 lAlaLysValLysGlnGlnGlnLysProAlaThrGlnThrGlnAlaGlnArgThrG 1045
Db 5272 AGGAGATTAAGCTTAAGAG-----GTATGGAACACATCTTCTTCCAA 5316
Qy 1045 ySerLysAlaLeuGlnLysValGlnGlnLysAlaGlnLysAspSerGlnLysAlaSerG 1065
Db 5317 TGCCAGCATGAAGAAAGAACTGAAAGGCTCAAAATGAGATGAGAAACCTTCTTAAGA 5376
Qy 1065 u-----LysGlnLysAspValMetProLysGlyProValGlnGlnAlaG 1080
Db 5377 GTTCACTCTTAAATGCTGAGAAAGACTCTTAAGTGA--GAGGTCAAGAT----- 5428
Qy 1080 yAlaGlnHisLeuAlaGlnGlnLysSerGlnLysGlnAlaThrProGluSerLeuGln 1100
Db 5429 -TTAAAGCATCAAGATGAAGAGTAATGTATCTTAACACAGTAACTTAAGGCGACAGAGA 5487

```


QY 1100 lProGluValThrAlaAspValAspHisValAlaThrCysGlnVal-----llely 1117
 Db 5488 ACATGATAACCAACGAATCTACTGAGAGGGAACACAGCTATACAGGTGAGACTGA 5547
 QY 1117 sLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSe 1137
 Db 5548 AGAGCAAGACTCTCTGAGTATGAGCACAGACCTTACATGTTTCAAGATCGTTCCATCAGC 5607
 QY 1137 rGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAs 1157
 Db 5608 GAAGAGT-----GCCAACCTCTGTGAAGTAAGGATTTTCAGCTCACATGATGAATAA 5661
 QY 1157 p-----GluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValarglnSe 1174
 Db 5662 TAATCTACTTACAGCAGATTTTACAGCTCAAGAAAGAAATGCTGGATAGAGGAGGAA 5721
 QY 1174 rGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAs 1194
 Db 5722 GCAGAAAAACAAGNATTTAGCCAGACTTTAGAAATGAGAAA---AATACCTTACTGAG 5778
 QY 1194 nAsnValProAlaGlnGluGlnHisGlyGluGluProGlyArgAspValLeuGluProTh 1214
 Db 5779 TCAGATATCAACAAGGAT-----GGTGAA-----CTAAAAATGCT 5814
 QY 1214 rGlnGlnGluLeuThr-----AlaAlaAlaValProva 1225
 Db 5815 TCAGGAGGAAGTAAACCAAAATGAACCTGTTAAATCAGCAATCAAGAAAGACTCTCCAG 5874
 QY 1225 lLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysVa 1245
 Db 5875 AGTTACCAAACTAAGGAGACACAGCAGAGAGAGAAAGATGATTTGGAAGAGAGGCTTAT 5934
 QY 1245 lLysGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspVa 1265
 Db 5935 GAATCAATTAGCAGAACTT-----AATCGAAGCATTGGGAATTTACTGTCAGGATGT 5985
 QY 1265 lThr-----TyrAspSerGluValMetGlyValAlaGlyCy 1277
 Db 5986 TACAGATGCCAAATAAAAAATGAGCTATTGGAACTCTGAATGAAGAACCTTAAAAAGTG 6045
 QY 1277 sGlnGluLysGluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluTh 1297
 Db 6046 TGTGAGTGAATTGGAGAGAAAGACGAGCTAGTTCAGAGAAAGAACTAAGGTGGAATC 6105
 QY 1297 rAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluG 1317
 Db 6106 AGAAATACGAAGAGATATTGCGAGAAATACAGGTGCTCAGAGAGAACCCGGAATAA 6165
 QY 1317 nGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMe 1337
 Db 6166 A-----AGCCATGCAAGAGAACTTCAGGAACTGTTTAA 6198
 QY 1337 tProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAs 1357
 Db 6199 AGAAAAACAACAGAGTAAGCAGCTA-----CAGAGGACTGCTATCAGGTA 6246
 QY 1357 pGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGl 1377
 Db 6247 TCAAGAGAAAAATTAGTCTCTGAGAGAACTGTTTAAAGCTCTAGAAATTTGTTCAAACTGA 6306
 QY 1377 nThrAlaGluAlaValGluLysValIleGluThrValValIleSerGluThrGlyGluse 1397
 Db 6307 ATCTCAAAAAGATTTTGGAAATAACCAAGAAATAA-----CTGGCTCAAGCAGTTGAACA 6360
 QY 1397 rProGluCysValGlyAlaHis-----LeuLeuProAlaGluLysSe 1411
 Db 6361 CCGCAAAAAGGCACAGCAAGTAATAGTACTTCAAAAGTCTCTAGATGACACTCAAG 6420
 QY 1411 rSerAlaThrGlyGlyHisThrProLeuGlnHisAlaGluAspThrValProLeuGlyPr 1431
 Db 6421 TGAAGCAGCA-----AGGTCCTAGCAGACAACTCAAGTTGAAAAA 6462
 QY 1431 oGluSerGlnAla-----GluSerIleProIleValThrProAlaProGluSerTh 1449

Db 6463 GGAACCTTCAGTCAATAAAGATTCAGTTAAAGCCGATGAACAAAGAGATGAA----- 6517
 QY 1449 rLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluG 1469
 Db 6518 -----GATCTTGACGGAAGACTGGAACAGCAGCAGAGAGAGAGACCTCTGAAGAGAA 6567
 QY 1469 uAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysVa 1489
 Db 6568 GAAGATATGCAAGAGAAACTGATGCTTTGGCAGAGAAAAAGTCCACTTGAAGAGAGAC 6627
 QY 1489 lLeu-----LysAlaGluProGluIleLeuGluLeuGluSe 1501
 Db 6628 AATGGAGAGATTCAGGTTACTTTGAACAAGAAAGAGAGAGGAGTTCAGCACTTCAGGA 6687
 QY 1501 rLysSerAsnLysIleVal 1507
 Db 6688 AAACCTGGACAGTACTGTG 6706
 RESULT 15
 US-09-724-676A-7240
 ; Sequence 7240, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 7240
 ; LENGTH: 8189
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-724-676A-7240
 Alignment Scores:
 Pred. No.: 1.98e-09 Length: 8189
 Score: 415.00 Matches: 316
 Percent Similarity: 38.17% Conservative: 305
 Best Local Similarity: 19.42% Mismatches: 669
 Query Match: 5.14% Indels: 337
 DB: 5 Gaps: 60
 US-09-902-432-4 (1-1596) x US-09-724-676A-7240 (1-8189)
 QY 26 LeuSerGly-----HisGlyProAlaAlaGluAlaSerGlyAlaAla----- 39
 Db 2404 CTCCTCAGGTGAAGGAACCTTAGCATGGTAACAGAAATTGAGGGCTCAGGTAAAGCAACTGGA 2463
 QY 40 GlyAspPro-AlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSe 59
 Db 2464 AATGAACCTTGCAAGACAGAGAGGCAAGAGACCTTGATTATGAAGC---CAAACTGC 2520
 QY 59 rSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGlnGluAsnGlnGluG 79
 Db 2521 CCATGACAAACCTGCTCCTCACTGAACAG-----ATCCATAGTCTCAGCATAGAACCAATC 2574
 QY 79 yGlnGluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgG 99
 Db 2575 TAAAGATGTGAATAAT-----GAAAGTTTACAGATGAAGTGGATGATGTGCAG-- 2623
 QY 99 uLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAs 119
 Db 2624 -----CTTCAGTTTCTCAGCAGAGTACCTTGATAAGAGAGCTCGCAAGGCCA 2670
 QY 119 pGlyGlnGluGluThrSerGluIleIle-----GluGlnIleProAlaSerGluAs 136
 Db 2671 GCTGCAAAATAGGAAAGTGAAGTGTCTGAGGGGCGAGAACGTCGTAAAGGCATATCTCAAG 2730
 QY 136 nAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPhelysLy 156
 Db 2731 TAAAGTGAAGAACTGTCCAGGCTCTTTTCAGAGAGGAAGAACTTGAATA----- 2779

QY 156 sValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGlnLysSerAs 176
 Db 2780 -----ACAAAAATGGAT----- 2791
 QY 176 pThrValGlnLeuLeuThrValLysLysAspGlnGlnGlnLysValAlaGlnAlaSerValG 196
 Db 2792 -----CAGCTCTTACTAGAGAAAAGAGATGTGGAAACCCCTCCAAACCAACCTCCA 2844
 QY 196 ValAlaGlnPheLysGlnGlnLysProSerValGlnThrAlaValGlnLysLeuSerLysG 216
 Db 2845 GGAGAGAGATCAACAAGTACAGAAATCAGCTTATGATGACTGAGAAAATGGTTCAGCT 2904
 QY 216 uSerGlnLeuLysGlnLysThrGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 236
 Db 2905 TATGAAAGAAATTTCTCTGGGGGTGAAATTAAGACTCTTAAAGAA----- 2953
 QY 236 rThrGlnLysProLeuGlnAlaGlnLysAspGlnAlaGlnGlnGlnLysAlaAspG 256
 Db 2954 -----CAGCTAAATTTTATTTCAGAGCTGAGAGCAAAAAAGACAGGTGAGAAAGA 3009
 QY 256 uGlnGlnGlnLysGlnGlnLysGlnLysProThrLysSerProGlnSerProSerProVa 276
 Db 3010 TATGAAAGTTCTCTGGCTTAAACAATAATATGATGATGATGATGATGATGATGATGAT 3069
 QY 276 LysSerGlnThrLysSerSerPheLysLysPheThrLysGlnLysLysLysLysLysLys 296
 Db 3070 AAGTAAAGAAAGAA-----CTTCAGCATGAAATTTGACCTTCTGAA 3108
 QY 296 GlyLysLysThrSerPheLysLysSerLys-----GlnLysAsp 308
 Db 3109 GAAAGAAATGAGCAGAGAAAGAAAGCTCCAGCAGCTTTATTAAACAGAAAGAGCT 3168
 QY 308 pLeuGlnThrAlaGlnLysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 328
 Db 3169 TCTGCAAGAGTCAAGTATGAGATGGAAGAAATTAAGCACTTGAAGATCAATCTAAGA 3228
 QY 328 sGlnLysThrGlnProLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 348
 Db 3229 AGAAATCCCACTCAGTGAAGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3288
 QY 348 gLeuSer-----AlaAspLysGlnLysValGlnLeuProLeuGlnLysAspG 363
 Db 3289 ATACTAGAAAAATGTGTGACTTCTAAGTCCAGAGAAATTAATTTATTAAACAGAC 3348
 QY 363 nValGlnAspLeuGlnAla-----SerSerGlnGlnLysCysAl 376
 Db 3349 AATATCTGAGAAAGAGTGAATCTACAGCATATAAGAAAGATTGGAAGAAAG----- 3403
 QY 376 aProLeuAlaThrGlnLysLysPheAspGlnLysMetGlnLysLysGlnGlnLysValAla 396
 Db 3404 -----CTGGCAGCTGAA-----GAGCAATTTCCAGGCTGTGCTCAAA 3441
 QY 396 uValHisValSerThrValGlnLysThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 416
 Db 3442 GATGAATTCAGACTTCCAAAGATTAACAAACCA----- 3475
 QY 416 uGlnGlnLysValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 436
 Db 3476 -----ATAGATTGCTCCAGCAAGAAATACGTGAAAA 3507
 QY 436 oGlnGlnLysProGlnGlnLysAlaGlnLysProLysLysSerLysSerLysMetCys 456
 Db 3508 CCAAGCAATTAATCCAGAG-----TTATCAACAAGTAACACGGA 3546
 QY 456 sValSerGlnLysAspHisThrGlnLysThrAsp-----LeuSerProGlnGln 472
 Db 3547 TGCAGACTGATGGGACTCCGTAGCACTTTAAAGAAACAGGTGATTAAGTCCACCTTG 3606
 QY 472 uLysThrLeuProLysLysProGlnGlnLysLysLysLysLysLysLysLysLysLys 492
 Db 3607 TACAGGTAGTAGTGAACCTGGAAACCAAGAACTAAGAAAGAAATACTGCGCTTGAAAA 3666

QY 492 nGlnLysGlnLysValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 512
 Db 3667 AGAAAAAGAGCAACTTCAAAAAAGAGCTACAGGAACCTTAACCTTCCGCAAGCAATTC 3726
 QY 512 sLysLeuSerGlnLysLysGlnLysGlnLysLysLysLysLysLysLysLysLysLys 532
 Db 3727 TAAAAAGCAGCAGAGAAAGAAAGAAAGCATCTCAG-----GAGAGCTTAA 3771
 QY 532 yGlnLysGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 552
 Db 3772 GCACAGAAAGATGACTTAATTCGCTTCAAGAACAGTTTATGATGAGCAACAGAA 3831
 QY 552 r-----SerAl 554
 Db 3832 TGAGAAATTTGAGACCAAGCTTAAGCAATCCAGATTCAAGTAAAGGAATCCATGACGG 3891
 QY 554 aSerSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 574
 Db 3892 AAAACTCCCAAGCAGACAGACAGAGAAATCGTGTCTTCCACTCAGGTTTAAAGAAAC 3951
 QY 574 oGlnAspGlnGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 584
 Db 3952 TTTATCAAGCCACAGAACAGCATCACACTCACTGTTTAAAGTCAACTTGCGCC 4011
 QY 585 -----SerAspGlnLysLysLysLysLysLysLysLysLysLysLysLysLys 599
 Db 4012 AGACTGCGCTTCTCACTTCTGAGATGCGAGTCTGCGAGGCGGAACTTCTGTTGCCA 4071
 QY 599 rPheLysLysMetValThrProLysLysArgValArgArgProSerGlnSerAspLysG 619
 Db 4072 GATTAAAGGCCAGCTG-----AAGGA 4092
 QY 619 uGlnGlnLysGlnLysValLys-----SerAlaThrLeuSerThrAspSerThrValSe 638
 Db 4093 AATTAAGCTGAGAAAGATTAAGTGAATGAAATGTTGTTCTCAACAAGAGACTTAA 4152
 QY 638 rGlnMetGlnAspGlnLysLysThrValGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 658
 Db 4153 TAAAAATTCAGAGAGGATTAATTCAGTTACAGAGAGATTAATAAACAAGGCTTAA 4212
 QY 658 gArgValAspThrSerValSerThrGlnLysLysLysLysLysLysLysLysLysLys 678
 Db 4213 CGAGAGCTTAAGACAGATATCCATGAAGCTGAAGTCCATGCCAAACCTCCACAGAA 4272
 QY 678 gAlaArgLysLysSerSerSerAspAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 698
 Db 4273 ATTG-----GAAAGCAGCCAACTACAATTTGCTGCGCTTGA 4308
 QY 698 rHisArgLysGlnGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLys 718
 Db 4309 ACATCTTAAGAGAAATTCGAACCTTAACCTGAGTGAACCTG-----CAAAACTCAT 4359
 QY 718 rThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 735
 Db 4360 CAAAAG 4419
 QY 735 ySerProSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 755
 Db 4420 CACTAAAAATACAGACAGATTAATGAACAAGAAATTTAATTAAGGCTGCTCATACA 4477
 QY 755 oArgLysLysSerLysSerLysLysLysLysLysLysLysLysLysLysLysLysLys 775
 Db 4478 -----CAGCTTAAGAAATCAAGCCAAAGAGATGATGAGAGATTAACAGCT 4524
 QY 775 uSerThrGlnLys-----GlnProSerArgGlnGlnGlnGlnGlnGlnGlnGlnGln 788
 Db 4525 ACAGGTGAACTTGTGAATTAAGCAAAAAACCAAGAGATGAGAGAGAGAGAGAGAG 4579
 QY 788 LserLysLysPheLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 808
 Db 4580 -----AGAGCAAAAGCAACAATTAACAAGAAATCTGCAAGC 4614
 QY 808 nAlaThrValGlnAspSerGlnProValGlnLysAsnGlnLysAspProAsnValProAl 828

[illegible]

Db	5608	GAAGAGT-----GCCAACCCCTGCTGTAGTAGAAGATTTCAGCTTCACACATGATGAATAATAA	5666
Qy	1157	p-----GluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSe	1174
Db	5662	TAACTACTACACGAGATTGTAGCTCAAGAGAAAGAAATTCCTGGATTAGAGGAGGAGAA	5721
Qy	1174	rGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAs	1194
Db	5722	GCAGAAAAACAAGGAATTTAGCCGACACTTTAGAAAATGAGAAA--AATACCTCTTACTGAG	5778
Qy	1194	naenValProAlaGlnGluHisGlnGluGluProGlyArgAspValLeuGluProTh	1214
Db	5779	TCAGATATCAACAAGAGAT-----GGTGAA-----CTAAAAATGCT	5814
Qy	1214	rGlnGlnGluLeuThr-----AlaAlaAlaValProVa	1225
Db	5815	TCAGGAGGAAGTAACCAAAATGAACCTGTTAAATCAGCAAAATCCAAGAGAAGACTCTCCAG	5874
Qy	1225	LeuAlaLysThrGluValGlnGlnGluGluValAspTrpLeuAspGlyGluLysVa	1245
Db	5875	AGTTTACCMAACTTAAGAGGACACGACGAAGACGAGAAAGATGATTTTGGGAAGAGAGCGCTTAT	5934
Qy	1245	LlysGlnGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspVa	1265
Db	5935	GAATCAATTAGCAGAACTT-----AATCGAAGCATTTGGGAATTACTCTCAGGATGT	5985
Qy	1265	lThr-----TyrAspSerGluValMetGlyValAlaGlyCy	1277
Db	5986	TACAGATGCCCAATAAAAAATCAGCTATTGGAATCTGAAATGAAGAACCTTAAAAAGTG	6045
Qy	1277	sGlnGluLysGluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluTh	1297
Db	6046	TGTGAGTGAATTTCGAAGAGAGAAAGCAGCAGTTAGTCAAGGAAAAAACTAAGGTGGGAATC	6105
Qy	1297	rAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGl	1317
Db	6106	AGAAATACGAAGAAGATATTTCGAGAAAAATACAGGTGCTCAGAAAGAACCCCGGAATAAA	6165
Qy	1317	nGluThrAlaAlaProGluHisGlnGlyThrTyrGlyLysProValLeuThrLeuAspMe	1337
Db	6166	A-----AGCCATCCAAAGGAACTTCAGGAAGCTTTAAA	6198
Qy	1337	tProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAs	1357
Db	6199	AGAAAAACAACAAGAGTAAAGCAGCTA-----CAGAAGGAGCTCATCAGGTA	6246
Qy	1357	pGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGl	1377
Db	6247	TCAAGAGAAAATTTAGTGCTCTGGAGAGAACTGTTAAAGCTCTAGAAATTTGTTCAAACATGA	6306
Qy	1377	nThrAlaGluAlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSe	1397
Db	6307	ATCTCAAAAAGATTTCGAATAATACCAAGAAAT-----CTGGCTCAAGCAGTTGAACA	6360
Qy	1397	rProGluCysValGlyAlaHis-----LeuLeuProAlaGluLysSe	1411
Db	6361	CCGCAAAAAGGCCACAAGCAGAAATTAGCTTCAAAGTCTCTGCTAGATGACACTCAAAG	6420
Qy	1411	rSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGlyPr	1431
Db	6421	TGAAGCAGCA-----AGGGTCCTAGCAGCAATCTCAAGTTTGAANA	6462
Qy	1431	oGluSerGlnAla-----GluSerIleProIleIleValThrProAlaProGluSerTh	1449
Db	6463	GGAACTTCAGTCAAAATAAGAAATCAGTTAAAAGCCAGATGAACAAAGGATGAA-----	6517
Qy	1449	rLeuHisProAspLeuGlnGlyIleGlySerAlaSerGlnArgGluArgSerGluGluGl	1469
Db	6518	-----GATCTTGACGAGACTGCGACAGCAGCAGAGAGACACTCGAAAGAGAA	6567
Qy	1469	uAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysVa	1489
Db	6568	GAAGAATATCGAAGAGAAAATCGATGCTTTTGGCGAGAGAAAAGTCCACTTTGGAAGAGAC	6627

```

Oy 1489 lleu-----LysAlaGluProGluIleLeuGluLeuGluSe 1501
      ::      ::      ::      ::      ::      ::      ::      ::
Db 6628 AATTGGAGAGATTCAAGTTACTTTGAACAAGAAAGCAAGATTCAAGCAACTTCAGGA 6687
Oy 1501 rlySerAsnLysIleVal 1507
      ::      ::      ::
Db 6688 AACTTGGACAGTACTGTG 6706

```

Search completed: December 13, 2002, 05:34:53
 Job time : 2025 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 12, 2002, 14:13:30 ; Search time 14973 Seconds
(without alignments)
11973.104 Million cell updates/sec
Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cgggcccggggagtagaag.....aaataaacattattgactac 6160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_cm:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	6160	100.0	6160	10	RNU23146	U23146 Rattus norv
2	5204.2	84.5	5236	10	RNU41453	U41453 Rattus norv
3	4831.4	78.4	5134	6	AR071282	AR071282 Sequence
4	4559.6	74.0	6197	10	AF326228	AF326228 Mus muscu
5	4558	74.0	6195	10	AB020886	AB020886 Mus muscu
6	4441.8	72.1	178954	2	AC115647	AC115647 Rattus no
7	4304	69.9	5902	10	AF326230	AF326230 Mus muscu
8	2179.6	35.4	6614	6	AX405857	AX405857 Sequence
9	2144.4	34.8	6608	6	AX329766	AX329766 Sequence
10	2144.4	34.8	6608	6	AX410793	AX410793 Sequence
11	2144.4	34.8	6608	9	HSU81607	UF01607 Homo sapien
12	2136.8	34.7	6297	9	AF001504	AF001504 Homo sapi
13	2118.8	34.4	6605	6	AR002603	AR002603 Sequence
14	2089.6	33.9	6287	9	AB003476	AB003476 Homo sapi
15	1906.4	30.9	104939	9	AL590413	AL590413 Human DNA
16	1906.4	30.9	171158	2	AC023201	AC023201 Homo sapi
17	923.8	15.0	102822	2	AC097731	AC097731 Rattus no
18	911.8	14.8	178954	2	AC115647	AC115647 Rattus no
19	681.2	11.1	129179	2	AC117205	AC117205 Mus muscu
20	651	10.6	102822	2	AC097731	AC097731 Rattus no
21	550	8.9	563	6	AX401547	AX401547 Sequence
22	466.4	7.6	1027	9	BC000188	BC000188 Homo sapi
23	466.4	7.6	1027	9	BC022814	BC022814 Homo sapi
24	416	6.8	8439	5	AF308810	AF308810 Xenopus l
25	313.2	5.1	1981	9	HUNGRAVIN	M963322 Human gravi
26	232.8	3.8	1058	11	G06673	G06673 human STS W
27	163.4	2.7	129179	2	AC117205	AC117205 Mus muscu
28	162.2	2.6	178	6	AX401799	AX401799 Sequence
29	162.2	2.6	178	10	RNU75404	U75404 Rattus norv
30	160.6	2.6	696	9	HUMZD39H03	AF086250 Homo sapi
31	131.8	2.1	386	11	G21790	G21790 human STS W
32	108.6	1.8	179145	9	AC009785	AC009785 Homo sapi
33	107.8	1.7	195690	10	AC098726	AC098726 Mus muscu
34	104.8	1.7	177681	2	AC116238	AC116238 Rattus no
35	104.6	1.7	174990	10	AC027653	AC027653 Mus muscu
36	101.6	1.6	35848	2	AC011553	AC011553 Homo sapi
37	101.2	1.6	33082	2	AC091365	AC091365 Rattus no
38	101	1.6	193363	2	AL627104	AL627104 Mus muscu
39	99.8	1.6	7218	6	I66494	I66494 Sequence 14
40	99.8	1.6	68100	2	AC105787	AC105787 Rattus no
41	99.6	1.6	196194	2	AC104920	AC104920 Mus muscu
42	99.2	1.6	206497	10	AC084389	AC084389 Mus muscu
43	98.8	1.6	45632	2	AC098158	AC098158 Rattus no
44	98.6	1.6	125816	10	AL591127	AL591127 Mouse DNA
45	98.2	1.6	194334	2	AC101527	AC101527 Mus muscu

ALIGNMENTS

RESULT 1
RNU23146
LOCUS Rattus norvegicus mitogenic regulation ssecks (322) gene, complete cds.
DEFINITION RNU23146
ACCESSION U23146
VERSION U23146.1 GI:1022682
KEYWORDS Rattus norvegicus.
SOURCE Rattus norvegicus
ORGANISM Rattus
REFERENCE 1 (bases 1 to 6160)
AUTHORS Lin,X., Nelson,P.J., Frankfort,B., Tomblor,E., Johnson,R. and

RNU23146 6160 bp DNA linear ROD 19-OCT-1995
Rattus norvegicus mitogenic regulation ssecks (322) gene, complete cds.

U23146
U23146.1 GI:1022682
Rattus norvegicus.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

TITLE Gelman, I.H.
Isolation and characterization of a novel mitogenic regulatory gene, 322, which is transcriptionally suppressed in cells transformed by src and ras

JOURNAL Mol. Cell. Biol. 15 (5), 2754-2762 (1995)

MEDLINE 95257957

PUBMED 7739556

REFERENCE 2 (bases 1 to 6160)

AUTHORS Gelman, I.H.

TITLE Direct Submission

JOURNAL Submitted (21-MAR-1995) Irwin H. Gelman, Microbiology, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029, USA

COMMENT On Oct 19, 1995 this sequence version replaced gi:733449.

FEATURES

source

1..6160

/organism="Rattus norvegicus"

/strain="fisher"

/db_xref="taxon:10116"

/clone="13.2.2"

/cell_line="Rat-6"

/cell_type="fibroblast"

/issue_type="embryo"

/clone_lib="RI-a (Robert Kraus, Mount Sinai School of Medicine)"

1..32

33..5044

/gene="322"

33..4823

/gene="322"

/function="possible mitogenic regulatory gene"

/note="Src Suppressed C Kinase Substrate."

/citation=[1]

/codon_start=1

/evidence=experimental

/product="SseCKS"

/protein_id="AA79517.1"

/db_xref="GI:1022683"

/translation="MGAGSSTGERSPEOPAGSDTPSELVLSGHGPAABASGACDPAADADPATKLPQKNGOLSVNGVAEOGDVHVOENOGOEVEVDEVGORESEVDKORVEMAAASTAVEDITDKOQETSEIIEQIPASENNVGMOVPASQANDVPFKVFPFVGFPTVKDKNEKSDTVQLTVKKBEGEAEVAGDHDPEVETVAGSASESE LKOSTREOGLKOBOSSTETPLQASQOAEERAKDGEKREKPEKSPSPSV NSSETTSFKFTFGHMGWRKKTSPKSKEDDLTAERKQOEAKDEBEKTEBA SEOEPAEDTQALISADYEKVELPLEDQVGLTASSEKCAPLATEVEPKMAHOE VVAEVHSTVEKTEEBEOGGGEAGGVVEGTGSLPPEKLAEOVEQOEAPBEIM KSRMCVSGDHOTLDLSPEKTLPKHBEIVSEVEMLSQERIKVOGSPKLKLFSS SGLKIKGKOKKRGGGGDEPGEYOHITHESPESADDEOKGESSASSPEPEETCL EKGPLERPODEABEGTTSDEKKEKREGITTPASPKKWTTPKKVRPSSEKRELEK VKSATLSTSTSTVSEMODEVKTVEBEQPEEPKRRVDTSVSEWALICVGSKKARKA SSSDDEGPRTLGGDSHRAEASAKDEKEDSVEQLSTEIEPSESEVWSIKPFGRR EGVSTWESFKRLVTPRKSKSKLEKEDSVEQLSTEIEPSESEVWSIKPFGRR KKRADGKOEOATVEDSGVEINEDDPNVPVAVPSEVAVAREMEAGNTELPOLLG AVVYSELSKTLVHTVSVAVIDGTRAVTSVERSPMSISATVETPLEHTEGAMPVE EYTKDITAEETPVLTQTLPEKGAHDMNTSEVDTPSEATATFETSALRTEVTA SGAETTDMSVAVSQLTQSPDTEEATPVQVEGVLDEEETQTLILQAVDKK EESQVATQVORTGSKALEKVEVEEDSEVLAEKEDVMPKOPVQEAHEHAAQK EITGOATPESLPEPTADVAVTCOVAKILOLMEQAAVAPSSSTLPLNNSGSTPLA DSDTADTQOODETTDSODSKATAVROGSOVTEEAAQTGKEPESTPLNVPAGEHGE EPGRDVLEPTQOELTAAVPIYATKTEVQGEVUMLOEKKEKQOEYVHVSQPSOKA ADVIYDEVMVAGQOEKESTEVOSLSLEGEMETDVKEKRETKPEVQSEGEQETA APHEHGVGKPVLTLDMPSSERKALGSLGSPSLPDQDKAGCTLEVOQSLDTVTOT AEAEIKVETVVISETGSPCEVGAHLPAKKSATGQHWTLQHAEDVAPPGPSQAE SIIPIVPAPESTLHPDLOGEISASORRSEEDKPDGADKESIAIKVIAKAP EILFESKSNKIVINVIOTAVDOPARTETTABETHAYDSOQCRDLSRPNKWTM KMDAKMKHPVQPRREDLOVTLVLEMAQPRKCLRLOLKAIVSK"

misc_signal

1581..1590

/gene="322"

/note="encodes nuclear localization signal; NLS"

/citation=[1]

1797..1811

/gene="322"

/note="encodes nuclear localization signal; NLS"

misc_signal

1797..1811

/gene="322"

/note="encodes nuclear localization signal; NLS"

misc_signal

2061..2073

/gene="322"

/note="encodes nuclear localization signal; NLS"

misc_signal

2301..2313

/gene="322"

/note="encodes nuclear localization signal; NLS"

misc_binding

4230..4358

/gene="322"

/note="encodes zinc-finger domain"

3' UTR

4816..6160

5032..5044

/gene="322"

/note="encodes nuclear localization signal; NLS"

polyA_signal

5760..5765

polyA_signal

6136..6141

polyA_site

6160

BASE COUNT 1853 a 1409 c 1753 g 1145 t

ORIGIN

Query Match 100.0%; Score 6160; DB 10; Length 6160;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 6160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCGCGCGGAGTAAAGAGCCACTGAGCCATGCGCGGAGGAGGAGTTCACCGAGAGC 60

Db 1 CGGCGCGGAGTAAAGAGCCACTGAGCCATGCGCGGAGGAGGAGTTCACCGAGAGC 60

QY 61 GGAGCCCCGAGACCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120

Db 61 GGAGCCCCGAGACCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120

QY 121 GGGCCGAGCTGAAGCCTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180

Db 121 GGGCCGAGCTGAAGCCTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180

QY 181 AGCTCCACAGAAATGCGCAGCTGTCTGTCACAGCGCGTGAACCAAGAGATG 240

Db 181 AGCTCCACAGAAATGCGCAGCTGTCTGTCACAGCGCGTGAACCAAGAGATG 240

QY 241 TCCATGTCGAAGAGAAACCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

Db 241 TCCATGTCGAAGAGAAACCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

QY 301 GACGACGAGATCGAAGATGTGAGAGAAAGCCGAGTTGAAGAAATGCGGCCAAT 360

Db 301 GACGACGAGATCGAAGATGTGAGAGAAAGCCGAGTTGAAGAAATGCGGCCAAT 360

QY 361 CCACAGCTGTGAAGATATCAAAAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420

Db 361 CCACAGCTGTGAAGATATCAAAAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420

QY 421 AGATCCCTGCTTCAGAAAACAAATGTGAGAAATGTACAGCTGTAGTCCAGAGCTA 480

Db 421 AGATCCCTGCTTCAGAAAACAAATGTGAGAAATGTACAGCTGTAGTCCAGAGCTA 480

QY 481 ATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTCACGTTGAAGAG 540

Db 481 ATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTCACGTTGAAGAG 540

QY 541 ATAAATAAGAAAGTCAATGATCTGTCCACATCTGTCACAAAGAGTAAGGCGAAG 600

Db 541 ATAAATAAGAAAGTCAATGATCTGTCCACATCTGTCACAAAGAGTAAGGCGAAG 600

QY 601 GGGCAGAGCCTCTGTGCGAGCTGAGACCAACGAGAGCCAGTGTGAGAGCTCCGTG 660

Db 601 GGGCAGAGCCTCTGTGCGAGCTGAGACCAACGAGAGCCAGTGTGAGAGCTCCGTG 660

QY 661 GAGAGTCCAGATCCAAAGAAAGTGGAGCTGAGCAATCCACAGAGAAAGCAAGGACCC 720
Db 661 GAGAGTCCAGATCCAAAGAAAGTGGAGCTGAGCAATCCACAGAGAAAGCAAGGACCC 720
QY 721 TGAAGCAAGAAACAGAGCAGACAGAAATCCCCCTTCAAGCCGAATCTGATCAAGCGGCTG 780
Db 721 TGAAGCAAGAAACAGAGCAGACAGAAATCCCCCTTCAAGCCGAATCTGATCAAGCGGCTG 780
QY 781 AGGAAGCAAGCAAGATGAAGGAGAGAAAGAAACAAAGAGAAAGAGCCCAACCAAGTCCCCAG 840
Db 781 AGGAAGCAAGCAAGATGAAGGAGAGAAAGAAACAAAGAGAAAGAGCCCAACCAAGTCCCCAG 840
QY 841 AATCCCGAGAGCCCGCTGAGTCAACAGTGAAGCAATCTTCTTCAAGAGTCTTCACTC 900
Db 841 AATCCCGAGAGCCCGCTGAGTCAACAGTGAAGCAATCTTCTTCAAGAGTCTTCACTC 900
QY 901 ACGGTTGGGCGGCTGGCGCAAGAAAGACAGCTTCAAGAAATCAAAAGAGAGATCATCTGG 960
Db 901 ACGGTTGGGCGGCTGGCGCAAGAAAGACAGCTTCAAGAAATCAAAAGAGAGATCATCTGG 960
QY 961 AAATCGCGAGAGAAAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
Db 961 AAATCGCGAGAGAAAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
QY 1021 AGACAGAGCCAGCTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
Db 1021 AGACAGAGCCAGCTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
QY 1081 CAGCAGACTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Db 1081 CAGCAGACTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
QY 1141 CGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
Db 1141 CGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
QY 1201 ACCAAGAAAGTGTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
Db 1201 ACCAAGAAAGTGTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
QY 1261 GAGG 1320
Db 1261 GAGG 1320
QY 1321 CTGAGAAACTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
Db 1321 CTGAGAAACTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
QY 1381 TGAAGAGCAGAGAGATGTGTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1381 TGAAGAGCAGAGAGATGTGTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
QY 1441 CTGAAGAGAGAGCAGCTGCCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
Db 1441 CTGAAGAGAGAGCAGCTGCCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
QY 1501 CCTCTCAGGAAAGAACTCAAGGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
Db 1501 CCTCTCAGGAAAGAACTCAAGGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
QY 1561 GCTTAAAGAGAGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Db 1561 GCTTAAAGAGAGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
QY 1621 AGCCTGAGAGAGTACCAACAGATCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
Db 1621 AGCCTGAGAGAGTACCAACAGATCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
QY 1681 GAGAGAGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db 1681 GAGAGAGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
QY 1741 CGCTGGAAGCACCCAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800

Db 1741 CGCTGGAAGCACCCAGAGTGGGGAAGCTGAGGAAGAACTACTTCCGATGGAGAGAGA 1800
QY 1801 AGAGAGAAAGGAGTCACTCCCTGGGCACTCTTCAAAAAGATGGTGACACCCCAAGAAAACGGG 1860
Db 1801 AGAGAGAAAGGAGTCACTCCCTGGGCACTCTTCAAAAAGATGGTGACACCCCAAGAAAACGGG 1860
QY 1861 TCCGAAGAACCTTCTGAGAGTGAACAAGAGGAAGAGTGGAGAAAGGTCAAGAGGCCACCT 1920
Db 1861 TCCGAAGAACCTTCTGAGAGTGAACAAGAGGAAGAGTGGAGAAAGGTCAAGAGGCCACCT 1920
QY 1921 TGTCCTCCACTGATAGCACAGTGTCAAGATGAAGTGAAGTCAAAACTGTTGGTGAGG 1980
Db 1921 TGTCCTCCACTGATAGCACAGTGTCAAGATGAAGTGAAGTCAAAACTGTTGGTGAGG 1980
QY 1981 AACAAAGCCAGAGGAAACCAAGCGTAGGCTGATCTTCAAGTGTCTTGGGAAGCACTGA 2040
Db 1981 AACAAAGCCAGAGGAAACCAAGCGTAGGCTGATCTTCAAGTGTCTTGGGAAGCACTGA 2040
QY 2041 TTTGTGTCGATCATCCAAAGAGAGCAAGAAAGGAGCATCTTTCAGATGATGAAGGAG 2100
Db 2041 TTTGTGTCGATCATCCAAAGAGAGCAAGAAAGGAGCATCTTTCAGATGATGAAGGAG 2100
QY 2101 GGCCNAGGACACTGGGAGGGGACAGTCACAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Db 2101 GGCCNAGGACACTGGGAGGGGACAGTCACAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
QY 2161 CCGGAACAGACGCTGTTCTGCCAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Db 2161 CCGGAACAGACGCTGTTCTGCCAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
QY 2221 CACCCGAGCAGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Db 2221 CACCCGAGCAGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
QY 2281 AAAGATTAGTCACTCCNAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
Db 2281 AAAGATTAGTCACTCCNAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
QY 2341 CTAGTGTAGAGCAGTGTCTCCACTGAGATCGAACCGAGTGAAGAAATCTTGGGTTTCCA 2400
Db 2341 CTAGTGTAGAGCAGTGTCTCCACTGAGATCGAACCGAGTGAAGAAATCTTGGGTTTCCA 2400
QY 2401 TTAAGAAATTCATCCCCGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460
Db 2401 TTAAGAAATTCATCCCCGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460
QY 2461 CTGTGGAAGACTCAGGGCCAGTGGAGTAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520
Db 2461 CTGTGGAAGACTCAGGGCCAGTGGAGTAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520
QY 2521 TGCTCTGTCTGAGTATAATGCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2580
Db 2521 TGCTCTGTCTGAGTATAATGCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2580
QY 2581 AGCTGCCAGCTGCTGGGGCTGTGTACGCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2640
Db 2581 AGCTGCCAGCTGCTGGGGCTGTGTACGCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2640
QY 2641 ACCTGTGAGTGTGCGAGTCAATTTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2700
Db 2641 ACCTGTGAGTGTGCGAGTCAATTTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2700
QY 2701 CTCCTTCGTGATATCCGCTTCCGTTTCAAGAACTCTTTGAACACACAGCGGGAGGAGGAGGAGGAGGAGGAGGAGG 2760
Db 2701 CTCCTTCGTGATATCCGCTTCCGTTTCAAGAACTCTTTGAACACACAGCGGGAGGAGGAGGAGGAGGAGGAGGAGG 2760
QY 2761 TGCCACCTGTTGAAGAGGTCATCAAAAAGACATCATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2820
Db 2761 TGCCACCTGTTGAAGAGGTCATCAAAAAGACATCATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2820
QY 2821 CCGAGAGCTTACAGAGGTTAAAGATGCCCATGACGACATGGGTCAACGAGTGAAGTGAATTT 2880

Db 2821 CCCAGACGTTACAGAGGGTAAAGTGCCTATGACGACATGCTCACGATGAATGATTT 2880
Qy 2881 TCACCTCAGAAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTCTGAAGAAGTTA 2940
Db 2881 TCACCTCAGAAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTCTGAAGAAGTTA 2940
Qy 2941 CCGAAGCATCGGGGGCCGAAAGACACACAGACATGTGTCTCCGCACTTTCCAGCTGACTG 3000
Db 2941 CCGAAGCATCGGGGGCCGAAAGACACACAGACATGTGTCTCCGCACTTTCCAGCTGACTG 3000
Qy 3001 ACTCCCCAGACACACAGAGGAAGCCACCAGTTCAAGAGGTAGAGTGGTGTCTAG 3060
Db 3001 ACTCCCCAGACACACAGAGGAAGCCACCAGTTCAAGAGGTAGAGTGGTGTCTAG 3060
Qy 3061 ATACAGAAAGAGAGAGCCGACAGCGAGGCCATCTCCAAAGCCGTTGACAAAGTTGA 3120
Db 3061 ATACAGAAAGAGAGAGCCGACAGCGAGGCCATCTCCAAAGCCGTTGACAAAGTTGA 3120
Qy 3121 AAGAGAGATCCCAAGTGCCTGCAACCCAGACTGTGACAGAAACGGGGTCAAAAAGCACTGG 3180
Db 3121 AAGAGAGATCCCAAGTGCCTGCAACCCAGACTGTGACAGAAACGGGGTCAAAAAGCACTGG 3180
Qy 3181 AGAAGTTGAGAGAGGTAGAGAGGAGCTCCGAAAGTGCTGGCTTCGAGAAAGAGAGACG 3240
Db 3181 AGAAGTTGAGAGAGGTAGAGAGGAGCTCCGAAAGTGCTGGCTTCGAGAAAGAGAGACG 3240
Qy 3241 TTATGCCAAGAGAGCCCGTGCAGAAAGCTGAGCTGAGCATTTGACAGGGCTCTGAGA 3300
Db 3241 TTATGCCAAGAGAGCCCGTGCAGAAAGCTGAGCTGAGCATTTGACAGGGCTCTGAGA 3300
Qy 3301 CTGCAAGAGGCTACTCCAGAGAGCTTTGAAAGTCTCTGAAATGACGGCAATGTAGACATG 3360
Db 3301 CTGCAAGAGGCTACTCCAGAGAGCTTTGAAAGTCTCTGAAATGACGGCAATGTAGACATG 3360
Qy 3361 TCGCCAGCGCCAGGTTATCAAGCTCCAGAGCTGATGGAACAGGCCGCTGGCCCTGAGT 3420
Db 3361 TCGCCAGCGCCAGGTTATCAAGCTCCAGAGCTGATGGAACAGGCCGCTGGCCCTGAGT 3420
Qy 3421 CATCCGAAACCTTGAACAGACAGTGAACAATGGAAGCACTCCCTTACAGATTCAGACA 3480
Db 3421 CATCCGAAACCTTGAACAGACAGTGAACAATGGAAGCACTCCCTTACAGATTCAGACA 3480
Qy 3481 CTGCAAGATGAGACACAGCAAGATGAAACCATTTGACAGCCAGGACAGTAAAGCCACTGAG 3540
Db 3481 CTGCAAGATGAGACACAGCAAGATGAAACCATTTGACAGCCAGGACAGTAAAGCCACTGAG 3540
Qy 3541 CTGTAGGAGCATCAGAGTCAAGAAAGAGCGGCTACTGCTCAGAAAAGAGAGCCTT 3600
Db 3541 CTGTAGGAGCATCAGAGTCAAGAAAGAGCGGCTACTGCTCAGAAAAGAGAGCCTT 3600
Qy 3601 CGACACTACCTAATATGTTCCAGCCCAAGAAACAATGGGGAAGAACCAAGAAAGATG 3660
Db 3601 CGACACTACCTAATATGTTCCAGCCCAAGAAACAATGGGGAAGAACCAAGAAAGATG 3660
Qy 3661 TTCTTGAACCTACACAGCAAGAGCTTACTGCTGCAAGCCGCTCTTGGGAAAGACTG 3720
Db 3661 TTCTTGAACCTACACAGCAAGAGCTTACTGCTGCAAGCCGCTCTTGGGAAAGACTG 3720
Qy 3721 AGGTGGGTCAAGAGGCTGAGTTGACTGGTGGATGAGAAAAAGTCAAAAGAAAGACAG 3780
Db 3721 AGGTGGGTCAAGAGGCTGAGTTGACTGGTGGATGAGAAAAAGTCAAAAGAAAGACAG 3780
Qy 3781 AGGTGTTGTACCTCTGAGCCCAACAGTCAAAAGGCTGCTGATGTGACATATGACAGTG 3840
Db 3781 AGGTGTTGTACCTCTGAGCCCAACAGTCAAAAGGCTGCTGATGTGACATATGACAGTG 3840
Qy 3841 AAGTATGAGAGTGGCCGGGTGTCAAGAAAAAGAGATGCTAAAGTGCAGATCTTTAGCC 3900
Db 3841 AAGTATGAGAGTGGCCGGGTGTCAAGAAAAAGAGATGCTAAAGTGCAGATCTTTAGCC 3900
Qy 3901 TCGAGAGGAGAGATGGAACCTGACGTTGAAAGAGAAAAAGGAGACAAAGCCAGAGC 3960
Db 3901 TCGAGAGGAGAGATGGAACCTGACGTTGAAAGAGAAAAAGGAGACAAAGCCAGAGC 3960

Qy 3961 AAGTATGGAAGAAAGGTGAGACAGAAACAGCCGCTCTGAGCATGAAGAACTTACGGGA 4020
Db 3961 AAGTATGGAAGAAAGGTGAGACAGAAACAGCCGCTCTGAGCATGAAGAACTTACGGGA 4020
Qy 4021 AGCCAGTCTTGACACTTGAACATGCGCCAGCTCAGAGAGGGGAAAGGCACTGGAAAGCTTG 4080
Db 4021 AGCCAGTCTTGACACTTGAACATGCGCCAGCTCAGAGAGGGGAAAGGCACTGGAAAGCTTG 4080
Qy 4081 GAGGAAGCCCTTCTCTCCAGACCAAGCAAAAGCAGGTTGCATAGAGTTCAAGTTCAA 4140
Db 4081 GAGGAAGCCCTTCTCTCCAGACCAAGCAAAAGCAGGTTGCATAGAGTTCAAGTTCAA 4140
Qy 4141 GCTTGAACACAAAGTACTTCAACACAGAGCTGTGAAAAAGTCAATAGAAACGGTTG 4200
Db 4141 GCTTGAACACAAAGTACTTCAACACAGAGCTGTGAAAAAGTCAATAGAAACGGTTG 4200
Qy 4201 TGATTTGAGAGACAGGTGAAGTCCAGAGTGTGAGGTGACACTTATTAACAGCTGAGA 4260
Db 4201 TGATTTGAGAGACAGGTGAAGTCCAGAGTGTGAGGTGACACTTATTAACAGCTGAGA 4260
Qy 4261 AGTCCTCTGCAACGGGTGGCCACTGAACTCTTCAGATGACAGAGACAGGTACCCCTGG 4320
Db 4261 AGTCCTCTGCAACGGGTGGCCACTGAACTCTTCAGATGACAGAGACAGGTACCCCTGG 4320
Qy 4321 GGCTGAGCTCAGGAGCAATCCATCCCATCATATGTAATCTCTGCTCTGAAAGCACCC 4380
Db 4321 GGCTGAGCTCAGGAGCAATCCATCCCATCATATGTAATCTCTGCTCTGAAAGCACCC 4380
Qy 4381 TACATCTGACCTTCAAGAGGAATAAGCGCATCCAGAGAGACGATTCAGAGGAAGAG 4440
Db 4381 TACATCTGACCTTCAAGAGGAATAAGCGCATCCAGAGAGACGATTCAGAGGAAGAG 4440
Qy 4441 ACAAGCAGATGCTGTCTCTGATGCTGACGCAAGAGAGATTCAGCAATCGAAAAAGTCC 4500
Db 4441 ACAAGCAGATGCTGTCTCTGATGCTGACGCAAGAGAGATTCAGCAATCGAAAAAGTCC 4500
Qy 4501 TCAAGGCTGAACCTGAGATCTTGAACCTTGAAGATTAAGAGCAACAAATGTGCTGAAGC 4560
Db 4501 TCAAGGCTGAACCTGAGATCTTGAACCTTGAAGATTAAGAGCAACAAATGTGCTGAAGC 4560
Qy 4561 TCATTGAGACAGCGGTTGACAGATTCGACGTAAGAAACAGCCCGGAAACTCATGCTT 4620
Db 4561 TCATTGAGACAGCGGTTGAGCAAGTTCGACGTAAGAAACAGCCCGGAAACTCATGCTT 4620
Qy 4621 ATGATTCACAGACCCAGGTTCTTGATGCAAGGCTTGAACAGAGGAGCCCAACAGATGCT 4680
Db 4621 ATGATTCACAGACCCAGGTTCTTGATGCAAGGCTTGAACAGAGGAGCCCAACAGATGCT 4680
Qy 4681 GGACAAAATTAAGATATGCAAGTGAACACCCAGATGCGCACCCCAAGAGAGACTTGC 4740
Db 4681 GGACAAAATTAAGATATGCAAGTGAACACCCAGATGCGCACCCCAAGAGAGACTTGC 4740
Qy 4741 AAGTCTGACCGTTCTGGAAGCATGGGCTCAGCTCGGAAATGCTTGCCGGCTTGCAGT 4800
Db 4741 AAGTCTGACCGTTCTGGAAGCATGGGCTCAGCTCGGAAATGCTTGCCGGCTTGCAGT 4800
Qy 4801 TGAAGCGCCGGGTCAAAGTAACTTGAAGCTGCTCTTCAACCCAAAGATTCAAA 4860
Db 4801 TGAAGCGCCGGGTCAAAGTAACTTGAAGCTGCTCTTCAACCCAAAGATTCAAA 4860
Qy 4861 AGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTAAGCCAGAGAGCCAGTGTCT 4920
Db 4861 AGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTAAGCCAGAGAGCCAGTGTCT 4920
Qy 4921 CTGGAACCTTAACCAAGAAATCCCAAGCAACCAAGCCAAAGCTTAACGAGAGGGGCG 4980
Db 4921 CTGGAACCTTAACCAAGAAATCCCAAGCAACCAAGCCAAAGCTTAACGAGAGGGGCG 4980
Qy 4981 ATCCCCAAAAGTTGAGTTCAGAGAAAGAAATGTCTAACCAAGTCAATCAAGAGACA 5040
Db 4981 ATCCCCAAAAGTTGAGTTCAGAGAAAGAAATGTCTAACCAAGTCAATCAAGAGACA 5040

[illegible][illegible]

Query Match 84.5%; Score 5204.2; DB 10; Length 5236;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 5223; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 1 CGGGCCGGCGGAGTAGAAGCACTGAGCCATGGCGCGAGGAGTTCCACCGAGCAGC 60
Db 1 CGGGCCGGCGGAGTAGAAGCACTGAGCCATGGCGCGAGGAGTTCCACCGAGCAGC 60
QY 61 GGAGCCCCGAGACCCCGCGGAGCCGACGCGCGAGGAGCTGTGCTCAGTGGCCATG 120
Db 61 GGAGCCCCGAGACCCCGCGGAGCCGACGCGCGAGGAGCTGTGCTCAGTGGCCATG 120
QY 121 GGCCCGCAGCTGAAGCTCCGAGAGCTGAGACCCCGCCGACGCGAGCCCGCCACCA 180
Db 121 GGCCCGCAGCTGAAGCTCCGAGAGCTGAGACCCCGCCGACGCGAGCCCGCCACCA 180
QY 181 AGCTCCCAAGAAATGCGCAGCTGTCTTCTGTCACGCGGTAGCTGAACAAGAGATG 240
Db 181 AGCTCCCAAGAAATGCGCAGCTGTCTTCTGTCACGCGGTAGCTGAACAAGAGATG 240
QY 241 TTCATGTCGAAGAGAAACCAAGAGGGGCGAGAGAAAGTGGTGTATGAGATGTTG 300
Db 241 TTCATGTCGAAGAGAAACCAAGAGGGGCGAGAGAAAGTGGTGTATGAGATGTTG 300
QY 301 GACGCGGAGAGTCAGAAATGAGAGAAAAAGCCGAGTTGAAGAAATGGCGCCAACT 360
Db 301 GACGCGGAGAGTCAGAAATGAGAGAAAAAGCCGAGTTGAAGAAATGGCGCCAACT 360
QY 361 CCACAGCTGTTGAAGATATCAAAAGATGGCGAGAGAGACATCAAGAAATATTGAAC 420
Db 361 CCACAGCTGTTGAAGATATCAAAAGATGGCGAGAGAGACATCAAGAAATATTGAAC 420
QY 421 AGATCCCTGCTTCAGAAAACAATGTGGAAGAAATGTTACAGCTCTGATGCCAGGCTA 480
Db 421 AGATCCCTGCTTCAGAAAACAATGTGGAAGAAATGTTACAGCTCTGATGCCAGGCTA 480
QY 481 ATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTTAAATTCAGCTGAAGAGG 540
Db 481 ATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTTAAATTCAGCTGAAGAGG 540
QY 541 ATAAAAATGAAAAAGTCAGATACTGTCCAACTATCTCAAGAGATGAAGGCGAAG 600
Db 541 ATAAAAATGAAAAAGTCAGATACTGTCCAACTATCTCAAGAGATGAAGGCGAAG 600
QY 601 GGGCGAAAGCTCTGTGCGAGCTGAGAGCACAGAGGCCAGTGTGAGACTGCGCTG 660
Db 601 GGGCGAAAGCTCTGTGCGAGCTGAGAGCACAGAGGCCAGTGTGAGACTGCGCTG 660
QY 661 GAGAGTCAGCATCCAAGAAAGTGAAGCTGAAGCAATCCACAGAGAAAGAGGACCC 720
Db 661 GAGAGTCAGCATCCAAGAAAGTGAAGCTGAAGCAATCCACAGAGAAAGAGGACCC 720
QY 721 TGAAGCAAGAACAGACAGACAGAAATCCCTTCAAGCCGAATCTGATCAAGCGGCTG 780
Db 721 TGAAGCAAGAACAGACAGACAGAAATCCCTTCAAGCCGAATCTGATCAAGCGGCTG 780
QY 781 AGGAAGAGCCCAAAGATGTAAGGAAGAAAAACAAGAAAAAGCCCAAGTCCCGCAG 840
Db 781 AGGAAGAGCCCAAAGATGTAAGGAAGAAAAACAAGAAAAAGCCCAAGTCCCGCAG 840
QY 841 AATCCCGCAGAGCCCGCAGCAACAGTGAACAATCTTCTCAAGAGATTCTTCACTC 900
Db 841 AATCCCGCAGAGCCCGCAGCAACAGTGAACAATCTTCTCAAGAGATTCTTCACTC 900
QY 901 ACGGTTGGCGCGCTGGCCCAAGAAAGCAGTTCAAGAAATCAAAAAGAGATGATCTG 960
Db 901 ACGGTTGGCGCGCTGGCCCAAGAAAGCAGTTCAAGAAATCAAAAAGAGATGATCTG 960
QY 961 AAACTGCCAGAAAGAGAGAGCAAGAGCGAGCAAAAAAGTAAGCGAGAAAGAAAGAAA 1020
Db 961 AAACTGCCAGAAAGAGAGAGCAAGAGCGAGCAAAAAAGTAAGCGAGAAAGAAAGAAA 1020
QY 1021 AAGACGAGCAGCTCGAGAGAGCAGAGCGGCGAGAAAGACAGACAGACGAGGCTTGT 1080
Db 1021 AAGACGAGCAGCTCGAGAGAGCAGAGCGGCGAGAAAGACAGACAGACGAGGCTTGT 1080

QY 1081 CAGCAGACTAGAAAGGTGAGAGCTGCTTGGAAAGACAGATTGCTGACCTGAGGCAAT 1140
Db 1081 CAGCAGACTAGAAAGGTGAGAGCTGCTTGGAAAGACAGATTGCTGAGGCAAT 1140
QY 1141 CGTCAGAGAGAAAGTGTGCTCTTTGGCAAAGAAAGTTTGAATGAGAAATGGAACCC 1200
Db 1141 CGTCAGAGAGAAAGTGTGCTCTTTGGCAAAGAAAGTTTGAATGAGAAATGGAACCC 1200
QY 1201 ACCAAGAAAGTTGTTGCAAGAGTCCACCTGAGACCCGTGAGAAAGACAGAGAGAGG 1260
Db 1201 ACCAAGAAAGTTGTTGCAAGAGTCCACCTGAGACCCGTGAGAAAGACAGAGAGAGG 1260
QY 1261 GAGGAGGAGAGAGGCTGAAGAGGGGCGTGTGTGTAAGAAACAGAGAAATCTTGGCCC 1320
Db 1261 GAGGAGGAGAGAGGCTGAAGAGGGGCGTGTGTGTAAGAAACAGAGAAATCTTGGCCC 1320
QY 1321 CTGAAGAACTGGCTGAGCCCGCAGAGAGTCCCGCAGAAAGCTGAGCTGAGAGCTGA 1380
Db 1321 CTGAAGAACTGGCTGAGCCCGCAGAGAGTCCCGCAGAAAGCTGAGCTGAGAGCTGA 1380
QY 1381 TGAAGACAGAGATGTGTCTCTGAGAGAGACACACTCACTGACAGACCTTAAGTC 1440
Db 1381 TGAAGACAGAGATGTGTCTCTGAGAGAGACACACTCACTGACAGACCTTAAGTC 1440
QY 1441 CTGAAGAGAAAGACGCTGCCCAAACACCCAGAAAGCATTTGTCAGTAGAGTGAATGCTGT 1500
Db 1441 CTGAAGAGAAAGACGCTGCCCAAACACCCAGAAAGCATTTGTCAGTAGAGTGAATGCTGT 1500
QY 1501 CCTCTCAGGAAAGATCAAGGTACAGGAAAGTCCCTTGAAGAACTCTTCAGTACCTAG 1560
Db 1501 CCTCTCAGGAAAGATCAAGGTACAGGAAAGTCCCTTGAAGAACTCTTCAGTACCTAG 1560
QY 1561 GCTTAAAGAGCTGTCTGGAAAGAGCAAGAGGGGAAACGAGAGAGTGGGGAACCAAG 1620
Db 1561 GCTTAAAGAGCTGTCTGGAAAGAGCAAGAGGGGAAACGAGAGAGTGGGGAACCAAG 1620
QY 1621 AGCTGAGAAATACCAACACTTCAACCGAATCCCGAGAGTGTGATGAGAGAAAG 1680
Db 1621 AGCTGAGAAATACCAACACTTCAACCGAATCCCGAGAGTGTGATGAGAGAAAG 1680
QY 1681 GAGAGAGCTGTGCGTGTGCTCCCGAGAGAGCTGAGAGACCAAGTGTGAGAAAGGCG 1740
Db 1681 GAGAGAGCTGTGCGTGTGCTCCCGAGAGAGCTGAGAGACCAAGTGTGAGAAAGGCG 1740
QY 1741 CGCTGAGAGCAACCCAGAGATGGGGAAGCTGAGAGAGAACTTCTTCCATGGAAGAA 1800
Db 1741 CGCTGAGAGCAACCCAGAGATGGGGAAGCTGAGAGAGAACTTCTTCCATGGAAGAA 1800
QY 1801 AGAGAGAGGGATCACTCCCTGGGATCTTCAAAAAGATGAGTCAACCAAGAAACGGG 1860
Db 1801 AGAGAGAGGGATCACTCCCTGGGATCTTCAAAAAGATGAGTCAACCAAGAAACGGG 1860
QY 1861 TCCGAAGACCTTCTGAGAGTGAACAAGAGAGAGCTGGAAGAGTCAAGAGCGCACCT 1920
Db 1861 TCCGAAGACCTTCTGAGAGTGAACAAGAGAGAGCTGGAAGAGTCAAGAGCGCACCT 1920
QY 1921 TGTCTTCACTGATGCAACAGTGTAGAAATGCAAGATGAAGTCAAACTGTTGGTGAAG 1980
Db 1921 TGTCTTCACTGATGCAACAGTGTAGAAATGCAAGATGAAGTCAAACTGTTGGTGAAG 1980
QY 1981 AACAAAGCCAGAGAGCAAAAGCTAGGGTGAATCTTCAAGTGTGGAAGACCTGA 2040
Db 1981 AACAAAGCCAGAGAGCAAAAGCTAGGGTGAATCTTCAAGTGTGGAAGACCTGA 2040
QY 2041 TTTGTGTCGATCATCAAGAGAGAGCAAGAGGCAATCTTCAAGTGAAGAGG 2100
Db 2041 TTTGTGTCGATCATCAAGAGAGAGCAAGAGGCAATCTTCAAGTGAAGAGG 2100
QY 2101 GGCCAAAGCACTGGAGGGGCACTCACAGCAGAGAGAGGCGCAGCAAAAGCAAGAG 2160
Db 2101 GGCCAAAGCACTGGAGGGGCACTCACAGCAGAGAGAGGCGCAGCAAAAGCAAGAG 2160
QY 2161 CCGGAACAGACCTGTCTTCTGCAAGCAACCAAGAGCAAGGCAAGAGGTTCTT 2220

Db 2161 ||||| CCGGAAACAGAGCTGTTCTTCCAGGACACCCAGGAGCAGGACCAAGCCCAAGGAAGTTCTCT 2220
QY 2221 CACCCGAGCAGCGGGAGGCCCTTCCAGGGGAAAGGTGTCTCCACTTTGGAGTCAITTTA 2280
Db 2221 CACCCGAGCAGCGGGAGGCCCTTCCAGGGGAAAGGTGTCTCCACTTTGGAGTCAITTTA 2280
QY 2281 AAGATTAGTCACTCCCAAGAAAATAATCCAAAGTCAAACTGGGAAGAACCCGAGACT 2340
Db 2281 AAGATTAGTCACTCCCAAGAAAATAATCCAAAGTCAAACTGGGAAGAACCCGAGACT 2340
QY 2341 CTAGTGTAGAGCAGTGTCTCCACTGAGATCGAAACCGAGTGAAGAAATCTTTGGTTTCCA 2400
Db 2341 CTAGTGTAGAGCAGTGTCTCCACTGAGATCGAAACCGAGTGAAGAAATCTTTGGTTTCCA 2400
QY 2401 TTAAGAAATTCATCCCGGACCGGGGAAAGAAAGGGCAGACCGGAAGCAAGAACGCCA 2460
Db 2401 TTAAGAAATTCATCCCGGACCGGGGAAAGAAAGGGCAGATGGGAAGCAAGAACGCCA 2460
QY 2461 CTGTGGAAGACTCAGGCCAGTGGAGATAAATGAGGACGACCTTAATGTCCAGCCGTG 2520
Db 2461 CTGTGGAAGACTCAGGCCAGTGGAGATAAATGAGGACGACCTTAATGTCCAGCCGTG 2520
QY 2521 TGCCTCTGTCTGAGTATAATGCAAGTGAAGGAGAGAGATGGGAAGCCACGGGAATACGG 2580
Db 2521 TGCCTCTGTCTGAGTATGATGCAAGTGGAGGGGAGAGATGGGAAGCCACGGGAATACGG 2580
QY 2581 AGTGCCCCAGCTGCTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTC 2640
Db 2581 AGTGCCCCAGCTGCTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTC 2640
QY 2641 AACTGTGAGTGTGCGAGTCAITTTGATGGGACCAAGGCGAGTCAACAGTGTGCGAAGCGGT 2700
Db 2641 AACTGTGAGTGTGCGAGTCAITTTGATGGGACCAAGGCGAGTCAACAGTGTGCGAAGCGGT 2700
QY 2701 CTCTTCGTGATATCCGTTCCGTAAACAGACCTCTTTGAAACACACAGCGGGAGGCCA 2760
Db 2701 CTCTTCGTGATATCCGTTCCGTAAACAGACCTCTTTGAAACACACAGCGGGAGGCCA 2760
QY 2761 TGCCACCTGTTGAAGAGTCACTGMAAAGACATCATTTGAGAGAACTCCTGTGCTCA 2820
Db 2761 TGCCACCTGTTGAAGAGTCACTGMAAAGACATCATTTGAGAGAACTCCTGTGCTCA 2820
QY 2821 CCAGAGCTTACAGAGGTAAAGATGCCATGACGACATGGTCAACAGTGAAGTGAT 2880
Db 2821 CCAGAGCTTACAGAGGTAAAGATGCCATGACGACATGGTCAACAGTGAAGTGAT 2880
QY 2881 TCACCTCAGAAAGTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAAGTGA 2940
Db 2881 TCACCTCAGAAAGTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAAGTGA 2940
QY 2941 CCGAAGCATCGGGGGCCGAGAGACCAACAGATGTGTCCGAGTGTTCAGCTGACTG 3000
Db 2941 CCGAAGCATCGGGGGCCGAGAGACCAACAGATGTGTTCAGCTGACTG 3000
QY 3001 ACTCCAGACACACAGAGAAAGCCACCCAGTTACGAGGTAGAGAGTGTGTCTAG 3060
Db 3001 ACTCCAGACACACAGAGAAAGCCACCCAGTTACGAGGTAGAGAGTGTGTCTAG 3060
QY 3061 ATACAGAAAGAGGAGCGCCAGACCGAGCCATCTCCAGCGGTGACAGAGTGA 3120
Db 3061 ATACAGAAAGAGGAGCGCCAGACCGAGCCATCTCCAGCGGTGACAGAGTGA 3120
QY 3121 AAGAGGAGTCCCAGGTGCTTGCACCCAGACTGTGACAGAAACGGGTCAAAAGCACTGG 3180
Db 3121 AAGAGGAGTCCCAGGTGCTTGCACCCAGACTGTGACAGAAACGGGTCAAAAGCACTGG 3180
QY 3181 AGAAGGTTAGAGGTTAGAGGAGACTCCGAAGTGTGCTTCGAGAAAGAGAGACG 3240
Db 3181 AGAAGGTTAGAGGTTAGAGGAGACTCCGAAGTGTGCTTCGAGAAAGAGAGACG 3240
QY 3241 TTATGCCGAAGAACCCGTTCAGGAAGCTTGAGCTCTTGCACAGGGCTCTGAGA 3300

Db 3241 TTATGCCGAAGAACCCGTTCAGGAAGCTTGAGCTGAGCATCTTGACAGGGCTCTGAGA 3300
QY 3301 CTGGACAGGCTACTCCAGAGAGCCTTTGAAGTTCCTGAAGTCAACGGCAGATGTAGACCATG 3360
Db 3301 CTGGACAGGCTACTCCAGAGAGCCTTTGAAGTTCCTGAAGTCAACGGCAGATGTAGACCATG 3360
QY 3361 TCCGACAGTGCAGGTTATCAAGCTCCAGCAGCTGATGGAAACAGGCGTGGCCCTGAGT 3420
Db 3361 TCCGACAGTGCAGGTTATCAAGCTCCAGCAGCTGATGGAAACAGGCGTGGCCCTGAGT 3420
QY 3421 CATCCGAAACCTTTGACAGCAGTGAACAATGGAAGCACTCCCTTAGCAGATTCAGACA 3480
Db 3421 CATCCGAAACCTTTGACAGCAGTGAACAATGGAAGCACTCCCTTAGCAGATTCAGACA 3480
QY 3481 CTGAGATGGGACACAGCAAGATGAAACCAATTGACGCCAGGACAGTAAAGCCACTGCAG 3540
Db 3481 CTGAGATGGGACACAGCAAGATGAAACCAATTGACGCCAGGACAGTAAAGCCACTGCAG 3540
QY 3541 CTGTGAGGAGTCAAGGTCACAGATGAGAGAGAGGCGGTACTGCTCAGAAAGAGAGGCTT 3600
Db 3541 CTGTGAGGAGTCAAGGTCACAGATGAGAGAGAGGCGGTACTGCTCAGAAAGAGAGGCTT 3600
QY 3601 CGACACTTACCTAATAATGTTCCAGCCAGGAAGAAACATGGGGAAGAACCCAGGAAGATG 3660
Db 3601 CGACACTTACCTAATAATGTTCCAGCCAGGAAGAAACATGGGGAAGAACCCAGGAAGATG 3660
QY 3661 TTCTTGAACCTTACACAGCAAGAGCTTACTGCTGACGCCGTGCCCGTTCTGGCAAAAGCTG 3720
Db 3661 TTCTTGAACCTTACACAGCAAGAGCTTACTGCTGACGCCGTGCCCGTTCTGGCAAAAGCTG 3720
QY 3721 AGTGCGGTCAAGAGGTGAGGTTGACTGCTGACGCCGTGCCCGTTCTGGCAAAAGCTG 3780
Db 3721 AGTGCGGTCAAGAGGTGAGGTTGACTGCTGACGCCGTGCCCGTTCTGGCAAAAGCTG 3780
QY 3781 AGTGTTTGTGACTCTGACCCAAACAGTCAAAAGCTCTGATGTGACATATGACAGTG 3840
Db 3781 AGTGTTTGTGACTCTGACCCAAACAGTCAAAAGCTCTGATGTGACATATGACAGTG 3840
QY 3841 AAGTGATGGAGTGGCGGGTGTGAGAAAGGAGAGTACTGAAAGTGCAGAGCTTTAGCC 3900
Db 3841 AAGTGATGGAGTGGCGGGTGTGAGAAAGGAGAGTACTGAAAGTGCAGAGCTTTAGCC 3900
QY 3901 TGAGAGGAGAGATGGAACCTGACGTTGAAAGAGAAAGGAGAGACAAAGCCAGAGC 3960
Db 3901 TGAGAGGAGAGATGGAACCTGACGTTGAAAGAGAAAGGAGAGACAAAGCCAGAGC 3960
QY 3961 AAGTGAGTGAAGAGTGAAGGAAACAGCCGCTCTGAGCATGAAGAACCTTACGGGA 4020
Db 3961 AAGTGAGTGAAGAGTGAAGGAAACAGCCGCTCTGAGCATGAAGAACCTTACGGGA 4020
QY 4021 AGCCAGTCTCTGACACTTGACATGCCAGCTCAGAGAGGGGAGGACCTGGGAAGCCTTG 4080
Db 4021 AGCCAGTCTCTGACACTTGACATGCCAGCTCAGAGAGGGGAGGACCTGGGAAGCCTTG 4080
QY 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGATAGAGTTCAAGTTCAAA 4140
Db 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGATAGAGTTCAAGTTCAAA 4140
QY 4141 GCCTGACACAAACAGTCACTCAAAACAGCAGAGCTGTGGAAGAGGTCATAGAACCGTTG 4200
Db 4141 GCCTGACACAAACAGTCACTCAAAACAGCAGAGCTGTGGAAGAGGTCATAGAACCGTTG 4200
QY 4201 TGATTTTCAGAGACAGGTGAAGTCCAGAGTGTGTAGTGTCACACTTTATTACAGCTGAGA 4260
Db 4201 TGATTTTCAGAGACAGGTGAAGTCCAGAGTGTGTAGTGTCACACTTTATTACAGCTGAGA 4260
QY 4261 AGTCTCTGCAACCGGTGGCCACTGGACTCTTTAGCATGACAGGACACCGGTACCCCTGG 4320
Db 4261 AGTCTCTGCAACCGGTGGCCACTGGACTCTTTAGCATGACAGGACACCGGTACCCCTGG 4320
QY 4321 GGCCTGAGTCTCAGGAGAGATCCATCCCAATCATAGTAACCTCTGCTCTGAAAGCACCC 4380
Db 4321 GGCCTGAGTCTCAGGAGAGATCCATCCCAATCATAGTAACCTCTGCTCTGAAAGCACCC 4380

QY 4381 TACATCTGACCTTCAAGAGAAATTAAGCCATCCCAAGAGAGGCTGATCAAGAGAGAGG 4440
 Db 4381 TACATCTGACCTTCAAGAGAAATTAAGCCATCCCAAGAGAGAGGCTGATCAAGAGAGAGG 4440
 QY 4441 ACAAGCCAGATGCTGATCTGATGCTGACGGCAAGAGAGATCAAGCAATCGAAAAAGTCC 4500
 Db 4441 ACAAGCCAGATGCTGATCTGATGCTGACGGCAAGAGAGATCAAGCAATCGAAAAAGTCC 4500
 QY 4501 TCAAGGCTGAAGCTGAGATCTCTGAACTTGAAGATGAAGCAACAAGATTGTCTGAAAG 4560
 Db 4501 TCAAGGCTGAAGCTGAGATCTCTGAACTTGAAGATGAAGCAACAAGATTGTCTGAAAG 4560
 QY 4561 TCATTCAGACAGCCCTTGAACAGTTTGCACGTACAGAAACAGCCCCCGAAACTCATGCTT 4620
 Db 4561 TCATTCAGACAGCCCTTGAACAGTTTGCACGTACAGAAACAGCCCCCGAAACTCATGCTT 4620
 QY 4621 ATGATTCACAGACCCAGGTTCTGATGCAAGCTTGAAGAGAGAGAGCCCAACAGATGCT 4680
 Db 4621 ATGATTCACAGACCCAGGTTCTGATGCAAGCTTGAAGAGAGAGAGCCCAACAGATGCT 4680
 QY 4681 GGACAAATGAAGATGCTCAAGATGAAACACCCAGTGCAGCCGAGCCCAAGAGAGAGCTTGC 4740
 Db 4681 GGACAAATGAAGATGCTCAAGATGAAACACCCAGTGCAGCCGAGCCCAAGAGAGAGCTTGC 4740
 QY 4741 AAGTCTGACCGTTCTGAGAGGATGGCTCAGCTTCGAAATGCTTCCGCGCTTGCAGT 4800
 Db 4741 AAGTCTGACCGTTCTGAGAGGATGGCTCAGCTTCGAAATGCTTCCGCGCTTGCAGT 4800
 QY 4801 TGAAGGCGCGGTGTCAAGATGAAGATGAAGAGTCTCTCAACCCAAAGATCCAA 4860
 Db 4801 TGAAGGCGCGGTGTCAAGATGAAGATGAAGAGTCTCTCAACCCAAAGATCCAA 4860
 QY 4861 AGAGCATGCTGCTGATGCTCAGCTCCTCAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 4920
 Db 4861 AGAGCATGCTGCTGATGCTCAGCTCCTCAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 4920
 QY 4921 CTGGAACCTTAACCAAGATCTCAAGATCTCAAGATCTCAAGATCTCAAGATCTCAAGAT 4980
 Db 4921 CTGGAACCTTAACCAAGATCTCAAGATCTCAAGATCTCAAGATCTCAAGATCTCAAGAT 4980
 QY 4981 ATCCCAAAAGTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 5040
 Db 4981 ATCCCAAAAGTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 5040
 QY 5041 AGGCCAGGAGAGAGAGCTGACAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
 Db 5041 AGGCCAGGAGAGAGAGAGCTGACAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
 QY 5101 TTAGTGTCTCATTTGATCATCTGTGAAGACCAAGATGTGAAGACCAAGATGTGAAGATG 5160
 Db 5101 TTAGTGTCTCATTTGATCATCTGTGAAGACCAAGATGTGAAGACCAAGATGTGAAGATG 5160
 QY 5161 CTGCTGTTGGAGCTTGAAGACCAAGATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5220
 Db 5161 CTGCTGTTGGAGCTTGAAGACCAAGATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5220
 QY 5221 TCCATGATTTCCACC 5237
 Db 5221 TCCATGATTTCCACC 5236

RESULT 3
 LOCUS AR071282 5134 bp DNA linear PAT 18-FEB-2000
 DEFINITION Sequence 1 from patent US 5910442.
 ACCESSION AR071282
 VERSION AR071282.1 GI:7222170
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 5134)

AUTHORS Gelman, I.H.
 TITLE Tumor suppressor gene
 JOURNAL Patent: US 5910442-A 1 08-JUN-1999;
 FEATURES Location/Qualifiers
 source 1..5134
 /organism="unknown"
 BASE COUNT 1511 a 1167 c 1451 g 1005 t
 ORIGIN

Query Match 78.4%; Score 4831.4; DB 6; Length 5134;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 5048; Conservative 0; Mismatches 76; Indels 31; Gaps 14;

QY 1016 GGAAAGACAGAGCAGCTCTCGAG 1075
 Db 1 GGAAAGACAGAGCAGCTCTCGAG 60
 QY 1076 GTTGTACAGACACTACAG 1135
 Db 61 GTTGTACAGACACTACAG 120
 QY 1136 GGATTCGTACAG 1195
 Db 121 GGATTCGTACAG 180
 QY 1196 AGCCCACTCAAGAAATTTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1255
 Db 181 AGCCCACTCAAGAAATTTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 QY 1256 GCAGGAG 1315
 Db 241 GCAGGAG 300
 QY 1316 GCCCTTGAAGAACTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1375
 Db 301 GCCCTTGAAGAACTGTGTGAG 360
 QY 1376 GCTGATGAAGAGCAG 1435
 Db 361 GCTGATGAAGAGCAG 420
 QY 1436 AAGTCTTAAG 1495
 Db 421 AAGTCTTAAG 480
 QY 1496 GCTGTCTCTCAG 1555
 Db 481 GCTGTCTCTCAG 540
 QY 1556 CTCAGAGCTTAAG 1615
 Db 541 CTCAGAGCTTAAG 600
 QY 1616 CGAAGAGCTGAG 1675
 Db 601 CGAAGAGCTGAG 660
 QY 1676 GAAAGGAGAGAGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1735
 Db 661 GAAAGGAGAGAGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 720
 QY 1736 AGGGCCGCTGAG 1795
 Db 721 AGGGCCGCTGAG 775
 QY 1796 GAAGAAG 1855
 Db 776 GAAGAAG 835
 QY 1856 ACAGGTCGAG 1915
 Db 836 AC-GGTCCGAG 894

1916 CACCTTGTCTCTCCATGATAGCACAGTGTGAGAAATCGAAGATGAAGTCAAAATCTGTTGG 1975
Db |||||
895 CACCTTGTCTCTCCATGATAGCACAGTGTGAGAAATCGAAGATGAAGTCAAAATCTGTTGG 954
QY
1976 TGAGGAAACAAAGCCAGAGGAAACAAAGCGTGGGTGGATCTTTCAGTGTCTTGGGAAGC 2035
Db |||||
955 TGAGGAAACAAAGCCAGAGGAAACAAAGCGTGGGTGGATCTTTCAGTGTCTTGGGAAGC 1014
QY
2036 ACTGATTTTGTGTGGATCATCTCAAGAAAGAGAGAGCAAGGAGCATCTCTTCAGATGATGA 2095
Db |||||
1015 ACTGATTTTGTGTGGATCATCTCAAGAAAGAGAGCAAGGAGCATCTCTTCATCA--GATAT 1071
QY
2096 AGGAGGCCAAGAGACACTGGGA--GGGACAGTCAACAGAGAGAGGAGGCCAGCAAGAGACA 2154
Db |||||
1072 AAGAGGCCAAGAGACACTGGGAGGGGAGACAGTCAACAGAGAGAGGAGGCCAGCAAGAGACA 1131
QY
2155 AAGAGGCCAAGAGACAGTGTCTTCTCCAGACACCCAGAGAGAGGAGCAACAGCGCAAGGAA 2214
Db |||||
1132 AAGAGCC-GAAGAGCGTGTCTTCTCCAGAGACCCAGAGAGAGGAGCAACAGCGCAAGGAA 1190
QY
2215 GTTCTCACCCGAGCCAGCGGGAAGCCCTTCCGAAGGGGAAGGTGTCTCACTTGGGAGT 2274
Db |||||
1191 GTTCTCACCCGAGCCAGCGGGAAGCCCTTCCGAAGGGGAAGGTGTCTCCACTTGGGAGT 1250
QY
2275 CATTTAAAGATTAGTCACTCCAGAAABAAATCCAAAGTCAAAACTGGAAGAGAAAGCCG 2334
Db |||||
1251 CATTTAAAGATTAGTCACTCCAGAAABAAATCCAAAGTCAAAACTGGAAGAGAAAG 1310
QY
2335 AAGACTCTAGTGTAG-----AGCAGTTGTCCACTGAGATCGAACCGAGTAGAGAAGA 2386
Db |||||
1311 CCGAAGGACTCTAGTTGTAGGAGCAGGTTGTCCACTGAGATCGAACCGTTAGAGNAGA 1370
QY
2387 ATCTTGGGTTTCCATTAAGAAATTCATPCCCGGACCGCGGAAGAAAGGCGCAGACGGGAA 2446
Db |||||
1371 ATCTTGGGTTTCCATTAAGAAATTCATPCCCGGACCGCGGAAGAAAGGCGCAGATGGAA 1430
QY
2447 -GCAAGAACAGCCACTGTGGAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCTTA 2505
Db |||||
1431 GGCAAGAACAGCCACTGTGGAGACTCAGGGCCAGTGGAGATAAATGAGGAGCAGCCTG 1490
QY
2506 ATGTCCAGCGCTCGTGCCTCTGTCTGAGTATAATGAGTGGAGAGGAGAAAGATGGAAG 2565
Db |||||
1491 ATGTCCAGCGCTCGTGCCTCTGTCTGAGTATGATGAGTGGAGAGGAGAAAGATGGAAG 1550
QY
2566 CCAGGGGAATACGGAGCTGCCAGCTGCTGGGGCTGTGTACGTGTCCGAGGAGCTCA 2625
Db |||||
1551 CCAGGGGAATGCGGAGCTGCCAGCTGCTGGGGCTGTGT--AGTGTCCGAGGAGCTCA 1607
QY
2626 GTAAGACTCTGGTCCACACTGTGAGTGTCCAGTCAATTGATGGGACAGGGGAGTCACCA 2685
Db |||||
1608 GTAAGACTCTGGTCCACACTGTGAGTGTCCGAGTCAATTGATGGGACAGGGGAGTCACCA 1667
QY
2686 GTGTGGAAGAGCGGTCTCTTCTGGATATCCGCTTCCGTAACAGAACCTCTTGAACACA 2745
Db |||||
1668 GTGTGGAAGAGCGGTCTCTTCTGGATATCCGCTTCCGTAACAGAACCTCTTGAACACA 1727
QY
2746 CAGCGGAGAGCCATGCCACTGTGTGAAGAGTCACTGAAAGAGACATCAATTGCGAGAAG 2805
Db |||||
1728 CAGCGGAGAGCCATGCCACTGTGTGAAGAGTCACTGAAAGAGACATCAATTGCGAGAAG 1787
QY
2806 AACTCTGTGCTCACCCAGACCTTACCAGAGGTTAAGATGCCATGACGACATGGTCA 2865
Db |||||
1788 AACTCTGTGCTCACCCAGACCTTACCAGAGGTTAAGATGCCATGACGACATGGTCA 1847
QY
2866 CCAGTGAAGTGGATTTTCACTCTCAGAAGCTGTGACAGCCACAGAGACTCAGAGGCTCTCC 2925
Db |||||
1848 CCACTGAAGTGGATTTTCACTCTCAGAAGCTGTGACGCCACAGAGACTCAGAGGCTCTCC 1907
QY
2926 GTACTGAAGAAGTTACCGAAGCATTCGGGGCCGAAGAGACCAACAGACATGGTGTCCGCGAG 2985
Db |||||
1908 GTACTGAAGAAGTTACCGAAGCATTCGGGGCCGAAGAGACCAACAGACATGGTGTCCGCGAG 1967
QY
2986 TTTCCAGCTGACTGACTCCCCAGACACACACAGAGGAGCCACCCAGCTTCAGGAGGTAG 3045

1968 TTTCCAGCTGACTGACTCTCCAGACACACAGAGAAAGCCACCCAGTTCAGGAGGTAG 2027
QY |||||
3046 AGAGTGTGTGTCTAGATACAGAAAGAGAGGAGCCAGACGACGAGGACATCTCTCAAAGCCG 3105
Db |||||
2028 AGGTTGTGTGTCTAGATACAGAAAGAGAGGAGCCAGACGACGAGGACATCTCTCAAAGCCG 2087
QY |||||
3106 TTGCAGACAAGGTGAAAGAGGAGTCCAGGTGCTCTCAACCCAGACTGTGCAGAGAACGG 3165
Db |||||
2088 TTGCAGACAAGGTGAAAGAGGAGTCCAGGTGCTCTCAACCCAGACTGTGCAGAGAACGG 2147
QY |||||
3166 GGTCAAAAGCACTGGAGAGGTTGAGAGGTTAGAGAGGACTCCGAAGTGTCTGGCTTCGG 3225
Db |||||
2148 GGTCAAAAGCACTGGAGAGGTTGAGAGGTTAGAGAGGACTCCGAAGTGTCTGGCTTCGG 2207
QY |||||
3226 AGAAAGAGAAAGGACGTATTATGCCGAAAGGACCCGTGCAGGAAGCTGAGAGCTGAGCATCTTG 3285
Db |||||
2208 AGAAAGAGAAAGGACGTATTATGCCGAAAGGACCCGTGCAGGAAGCTGAGAGCTGAGCATCTTG 2267
QY |||||
3286 CACAGGGCTCTGAGACTGAGACAGGCTATCTCAGAGAGCCCTTGAAGTTCCTGAAAGTCAACGG 3345
Db |||||
2268 CACAGGGCTCTGAGACTGAGACAGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAAGTCAACAG 2327
QY |||||
3346 CAGATGTAGACCATCTCGCACGCTGCCAGGTTATCAAGCTCCAGAGCTGATGGAACAGG 3405
Db |||||
2328 CAGATGTAGACCATCTCGCACGCTGCCAGGTTATCAAGCTCCAGAGCTGATGGAACAGG 2387
QY |||||
3406 CCGTGCCCTCTGAGTCACTCCGAAACCTTGACAGACAGTGCAGACAAATGGAAGCACTCCCT 3465
Db |||||
2388 CCGTGCCCTCTGAGTCACTCCGAAACCTTGACAGACAGTGCAGACAAATGGAAGCACTCCCT 2447
QY |||||
3466 TAGCAGATTTCAGACACTGAGATGGGACACAGCAAGATGAAACCATTTGACAGCAGGACA 3525
Db |||||
2448 TAGCAGATTTCAGACACTGAGATGGGACACAGCAAGATGAAACCATTTGACAGCAGGACA 2507
QY |||||
3526 GTAAAGCCACTGAGCTGTGAGCAGTCAAGGTCAAGAGAGGCGGCTACTGTCTC 3585
Db |||||
2508 GTAAAGCCACTGAGCTGTGAGCAGTCAAGGTCAAGAGAGGCGGCTACTGTCTC 2567
QY |||||
3586 AGAAAGAGAGGCTTCCACACTACCTAAATGTTTCCAGCCAGGAAGAACATGGGGAAG 3645
Db |||||
2568 AGAAAGAGAGGCTTCCACACTACCTAAATGTTTCCAGCCAGGAAGAACATGGGGAAG 2627
QY |||||
3646 AACAGGAAGAGATGTTTCTGAACTTACACAGCAAGAGCTTACTGTGACGCGTCCCG 3705
Db |||||
2628 AACAGGAAGAGATGTTTCTGAACTTACACAGCAAGAGCTTACTGTGACGCGTCCCG 2687
QY |||||
3706 TTCTGCAAGAGTGAAGTGGGTCAAGAGGTTGAGGTTGACTGGTTGGATGGAGAAAAAG 3765
Db |||||
2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGTTGAGGTTGACTGGTTGGATGGAGAAAAAG 2747
QY |||||
3766 TCAAGAGAAACAGGAGGTGTTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGTGTATG 3825
Db |||||
2748 TCAAGAGAAACAGGAGGTGTTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGTGTATG 2807
QY |||||
3826 TGACATATGACAGTGAAGTGGAGTGGCGGCTGTGAGGAAAGGAGAGTACTGAAG 3885
Db |||||
2808 TGACATATGACAGTGAAGTGGAGTGGCGGCTGTGAGGAAAGGAGAGTACTGAAG 2867
QY |||||
3886 TGCAGAGTCTTAGCTTGGAGGAGGAGATGGAATACTGACGTTGAAAGGAGAAAAAGG 3945
Db |||||
2868 TGCAGAGTCTTAGCTTGGAGGAGGAGATGGAATACTGACGTTGAAAGGAGAAAAAGG 2927
QY |||||
3946 AGACAAAGCCAGAGCAAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTCTGAGCATG 4005
Db |||||
2928 AGACAAAGCCAGAGCAAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTCTGAGCATG 2987
QY |||||
4006 AAGGAACCTTACGGGAAGCCAGTCTTGCACACTTGCATGCTCCAGCTCAGAGAGGGGAAGG 4065
Db |||||
2988 AAGGAACCTTACGGGAAGCCAGTCTTGCACACTTGCATGCTCCAGCTCAGAGAGGGGAAGG 3047
QY |||||
4066 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGCAAGCAGGTTGATAG 4125

Db 3048 CACTGGGAGCCTTGGAGGAGCCCTTCTCTCCAGACCAAGCAAGAGTTGCATG 3107
Qy 4126 AGTTTCAAGTTCAAAAGCCTTGGAGCAACAAGTCACTCAAAAGAGAGAGCTGTGAAAAG 4185
Db 3108 AGGTTCAAGTTCAAAAGCCTTGGAGCAACAAGTCACTCAAAAGAGAGAGAGCTGTGAAAAG 3167
Qy 4186 TCATAGAAACGGTGTGTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGTGCACAT 4245
Db 3168 TCATAGAAACGGTGTGTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGTGCACAT 3227
Qy 4246 TATTAACCGCTGAGAGAGTCTCTGCAACGGGTGGCCACTGAGACTTTGACAGATGAGAG 4305
Db 3228 TATTAACCGCTGAGAGAGTCTCTGCAACGGGTGGCCACTGAGACTTTGACAGATGAGAG 3287
Qy 4306 ACACGGTACCCCTGGGGCCTGAGTCTCAGGCAAGATCCATCCATATAGTAATCTCTG 4365
Db 3288 ACACGGTACCCCTGGGGCCTGAGTCTCAGGCAAGATCCATCCATATAGTAATCTCTG 3347
Qy 4366 CTCTGAAAGCACTCTACATCTTGACCTAACAAGGAAATAGCGGATCCCAAGAGAGAGC 4425
Db 3348 CTCTGAAAGCACTCTACATCTTGACCTAACAAGGAAATAGCGGATCCCAAGAGAGAGC 3407
Qy 4426 GATCAG 4485
Db 3408 GATCAG 3467
Qy 4486 CAATGAAAAAGTCTCAAGAGCTGAACTGAGATCCTGGAATTTGAGAGTAAGAGACACA 4545
Db 3468 CAATGAAAAAGTCTCAAGAGCTGAACTGAGATCCTGGAATTTGAGAGTAAGAGACACA 3527
Qy 4546 AGATTGTGCTGAAGCTCATTTGACAGACCGGTTGACCAAGTCAAGAAACAGGCC 4605
Db 3528 AGATTGTGCTGAAGCTCATTTGACAGACCGGTTGACCAAGTCAAGAAACAGGCC 3587
Qy 4606 CCGAACTCATGCTTATGATTCACAGACCCAGGTTCTGCATGACGAGGCTTGAACAGAGG 4665
Db 3588 CCGAACTCATGCTTATGATTCACAGACCCAGGTTCTGCATGACGAGGCTTGAACAGAGG 3647
Qy 4666 AGCCCAACAGATGCTGACAAAGATGAAATGCAAGTGAACCAACAGAGTCCGACAGC 4725
Db 3648 AGCCCAACAGATGCTGACAAAGATGAAATGCAAGTGAACCAACAGAGTCCGACAGC 3707
Qy 4726 CCAGAGAGAGCTTGAAGTCTGACCGTTCTGAGAGCATGAGGCTCAGCTCCGAAATGCT 4785
Db 3708 CCAGAGAGAGCTTGAAGTCTGACCGTTCTGAGAGCATGAGGCTCAG-CTCGAAATGCT 3765
Qy 4786 TCCCGGCTTGAAGTGAAGAGCGGTTGCAAGTGAAGATTTGAGAGAGTGCCTCTCA 4845
Db 3766 TCCCGGCTTGAAGTGAAGAGCGGTTGCAAGTGAAGATTTGAGAGAGTGCCTCTCA 3825
Qy 4846 ACCCAAGATCCAAAGAGAGATGCTGATGAGCCCTCAGCTCCAAAGTTAGCCAGAG 4905
Db 3826 ACCCAAGATCCAAAGAGAGATGCTGATGAGCCCTCAGCTCCAAAGTTAGCCAGAG 3884
Qy 4906 CAGAGGCGAGTGTCTTGAAACCTTAACCAAGATCCCAAGACCAACAGAGAGAGAG 4965
Db 3885 CAGAGGCGAGTGTCTTGAAACCTTAACCAAGATCCCAAGACCAACAGAGAGAGAG 3941
Qy 4966 TAAACGAG 5025
Db 3942 TAAACGAG 4000
Qy 5026 CAGTCAAG 5085
Db 4001 CAGTCAAG 4060
Qy 5086 CAGAACTCTAAGATGTTAGTGTCTCATTTGATCTGTAAAGACCAAGATGTGAAAACAG 5145
Db 4061 CAGAACTCTAAGATGTTAGTGTCTCATTTGATCTGTAAAGACCAAGATGTGAAAACAG 4120
Qy 5146 TCACAG 5205
Db 4121 TCACAG 4179

Qy 5206 CAG 5265
Db 4180 CAG 4239
Qy 5266 TCATCGGAG 5325
Db 4240 TCATCGGAG 4299
Qy 5326 TTGATGCCGTCCTGTTATTTCTAAACATTTCTGCTGTTTCAAGAGAGAGAGAGAGAG 5385
Db 4300 TTGATGCCGTCCTGTTATTTCTAAACATTTCTGCTGTTTCAAGAGAGAGAGAGAGAG 4359
Qy 5386 TTGATGCCGTCCTGTTATTTCTAAACATTTCTGCTGTTTCAAGAGAGAGAGAGAGAG 5445
Db 4360 TTGATGCCGTCCTGTTATTTCTAAACATTTCTGCTGTTTCAAGAGAGAGAGAGAGAG 4419
Qy 5446 GCCATACCTAGTTCACATTCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5505
Db 4420 GCCATACCTAGTTCACATTCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4479
Qy 5506 TTATATGATGCTCTCTCTCTGTAATTTATTTATTTTCTTAAAGAGAGAGAGAGAGAG 5565
Db 4480 TTATATGATGCTCTCTCTCTGTAATTTATTTATTTTCTTAAAGAGAGAGAGAGAGAG 4539
Qy 5566 TTTTGTATTAAGCAAT 5625
Db 4540 TTTTGTATTAAGCAAT 4599
Qy 5626 GCGTCAACCTGATACCTGAG 5685
Db 4600 GCGTCAACCTGATACCTGAG 4659
Qy 5686 TTAAGTTTACTGATGCTTAATCTGAGGCTTGAATCTGAGGCTTGAATCTGAGGCTTGA 5745
Db 4660 TTAAGTTTACTGATGCTTAATCTGAGGCTTGAATCTGAGGCTTGAATCTGAGGCTTGA 4719
Qy 5746 TATGACAG 5805
Db 4720 TATGACAG 4779
Qy 5806 TACATGATTTTCTCCCT 5865
Db 4780 TACATGATTTTCTCCCT 4839
Qy 5866 GTGTGATTAAG 5925
Db 4840 GTGTGATTAAG 4899
Qy 5926 GATGTTTCTTAATGAG 5985
Db 4900 GATGTTTCTTAATGAG 4959
Qy 5986 GGACCGAATTCATATGAG 6045
Db 4960 GGACCGAATTCATATGAG 5019
Qy 6046 AATCTGAG 6105
Db 5020 AATCTGAG 5079
Qy 6106 TTCTCTTGAATTAATTTGATTTGAACAATGATTTGAATTAATTAATTAATTAATTAAT 6160
Db 5080 TTCTCTTGAATTAATTTGATTTGAACAATGATTTGAATTAATTAATTAATTAATTAAT 5134

RESULT 4
AF326228
LOCUS AF326228 6197 bp mRNA linear ROD 16-JUL-2001
DEFINITION Mus musculus PKC binding protein SSeCKS mRNA, complete cds.
ACCESSION AF326228
VERSION AF326228.1 GI:13236403
KEYWORDS

[illegible]

Db	1305	ACCAAGAAAGTTGTCAGAGAGTCACAGTGAAGACCGTGGAGAAAGTGAACGAAAGGCA--	1362
Qy	1261	GAGAGAGAGGAGAGGCTGAAGAGGGGGCGTGTGTAGAAAGAAACAGAGAAATCTTGGCCCC	1320
Db	1363	-AGGAGAGCAGAGGTGAGAGGGGATGTGTGTGTGAAAGATCCGGAGAAATCTTGGCCCC	1421
Qy	1321	CTGAGAAACTGGCTGAGCCCCCAGAGAGTCCCCCAGAAAGCTGAGCTTGAGAGAGCTTGA	1380
Db	1422	CTGAGAAACTGGCTGAGACCAGAGAGTCCCCCAGAAAGCTGAGCTTGAGAGAGCTTGA	1481
Qy	1381	TGAAGACAGAGATGTGTGTCTTGGAGAGAACCAACTCAATTGACAACTTAAGTC	1440
Db	1482	TGAAGACCAAGAGATGTGTGTCTTGGGGGGTGAACATCTCAGCTGCAATCTTAAGTC	1541
Qy	1441	CTGAAGAGAAAGAGCTGCCCCAAACCCCAAGAGGCAATTGTCAAGTAGAGTAGTGTCTGT	1500
Db	1542	CTGAAGAGAAAGATGTCTACCAAAACCCCGAGAGCAATTGTCAAGTAGAGTAGTGTCTGT	1601
Qy	1501	CCTCTCAGAGAAAGATCAAGGTACAGGGAAGTCCCTTGAAGAACTCTTCAAGTCTCAG	1560
Db	1602	CCTCTCAGAGAGAGATCAAGGTACAGGGAAGTCCCTTGAAGAACTCTTCAAGTCTCAG	1661
Qy	1561	GCTTAAAGAGCTGTCTGGGAGAGAGCAGAGGGGA--ACGAGAGAGTGGGGGAGAGC	1617
Db	1662	GCTTAAAGAGCTCTCCGGAGAGAGAGAGGGGAGAGAGAGAGAGAGCGCGGGGAGTG	1721
Qy	1618	AAGAGCCGAGAGATACCAACACTTCACACCGAATCCCCAGAGAGTGTGATGAGCAGA	1677
Db	1722	AAGAGCCGAGAGATACCAACACTTCACACCGAATCCCCAGAGAGTGTGATGAGCAGA	1781
Qy	1678	AGGAGAGAGCTCTGCGTGTGTCCCGCAGAGAGCCTGAGAGAACCACTGTCTGAGAAAG	1737
Db	1782	AGGAGAGAGCTCTGCGTGTGTCCCGTGAAGAGCCGAGAGAGATCGCGTGTCTGAGAAAG	1841
Qy	1738	GGCGCTGAGAGCAACCCAGAGATGGGAGAGCTGAGAGAACTACTTCCATGGAAGAGA	1797
Db	1842	GGCGCATCGGAAGCAACCCAGAGAGCGAAGCTGAGAGAGAGCACTTCCAGCGGAGAGA	1901
Qy	1798	AGAAAGAGAGGGGATCACTCCCGGGGCATCTTCAAAAAGATGGTGACACCAGAAAGC	1857
Db	1902	AGAAAGAGAGGGGATCACTCCCGGGGCATCTTCAAAAAGATGGTGACACCAGAAAGC	1961
Qy	1858	GGGTCCGAAGACTTCTTGAGAGTGAACAGAGAGAGAGAGCTGAGAAAGTCAAGAGCGCA	1917
Db	1962	GGGTCCGAAGACTTCTTGAGAGTGAACAGAGAGAGAGAGCTGAGAAAGTCAAGAGTCCCA	2021
Qy	1918	CTTGTCTCTCACTGATAGCAAGTGTCAAGAAATGCAAGTGAAGTCAAACTGTTGGTG	1977
Db	2022	CTTGTCTCTCACTGATAGCAAGTGTCTGGAATGCAAGTGAAGTCAAGAGCGTGTGGCG	2081
Qy	1978	AGGAACAAAGCCAGAGAGAACCAAGAGGTAGGGTGAATCTTCAGTGTCTTGGAGAGAC	2037
Db	2082	AGGAACAAAGGTGAGAGAGCCAAAGGCGCAGGGTGAATCTTCAAGTGTCTTGGAGAGCT	2141
Qy	2038	TGATTTGTGTGGGATCATCAAGAGAGACCAAGAGAGGATCTTTCAGATGATGAAG	2097
Db	2142	TGATTTGTGTGGATCTGTCCAGAGAGAGACGAGAGAGAGATCTTTCAGATGATGAAG	2201
Qy	2098	GAGGGCCAAAGACATGSGAGAGGGGACAGTCAAGAGCAGAGGAGCGCCAGCAAAAGCAAG	2157
Db	2202	GAGGGCCAAAGACATGSGAGAGGGGATGGCCACAGAGCGGAGGCGTTAGCAAAAGCAAG	2261
Qy	2158	AAGCCGGAACAGAGCTTCTTCTCCAGACCCAGAGAGCAGACCAAGGCGCAAGAGAGTT	2217
Db	2262	AAG-----CAAGTGTCTTCTTCTCCAGACCCAGAGACCAAGGCGCAAGAGAGTT	2315
Qy	2218	CCTCACCCGAGCAGCGGGAAGCCCTTCCAGAGGGGAGAGGTGTCTCACTTGGAGATCAT	2277
Db	2316	CCTCACCCGAGCAGCTGGAAGCCCTTCTGAAGGGAGAGGCGTCTCACTTGGAGATCAT	2375
Qy	2278	TTTAAAGATTAGTCACTCAAGAAAAAATCAAGTCAAACTGGAAGAGAAAGCCGAG	2337

Db	2376	TTTAGAGATTAGTGA	CTCCACGAAAAAATTC	CAAGTC	CAAAACTGG	AAGAGACCGG	2435
Qy	2338	ACTCTAGTGA	AGCAGTGTCC	ACTGAATCG	AACCGAGTAG	AGAAAGATCTTTGGGTTT	2397
Db	2436	ACTCCGGTGC	AGACAGATTGGCTCTCC	AGATCG	AATCCAACTG	AGAGGAATCTTGGGTTT	2495
Qy	2338	CCATTAA	GAATTTATCC	CCCGGAGGGGG	GAAGAAAGGGG	AGACCGGACCAAGAAC	2457
Db	2496	CCATTAA	GAATTTATTTCT	CTGGAGGGG	GAAGAAAGGGG	AGATGGGAGCCAGAAC	2555
Qy	2458	CCACTGTGA	AGACCTCAGGGCC	AGTGAAGAT	TAATTGAG	AGACCTCTAAATGTC	2517
Db	2556	CCGCGGTT	GAAGACTCCGGGGCC	AGAGAGAT	CATATGAG	AGACCCCGACGTC	2615
Qy	2518	TCGTGCTCTGTCT	GTAGTATATAT	CACTGAGAG	GGGAGGAAGTGA	-----AGCCC	2568
Db	2616	TTGTGCTCTGTCT	GTAGTATATAT	CACTGAGAG	GGGAGGAAGTGA	-----AGCCC	2675
Qy	2569	AGGGGA	ATACGAGCTGCCCC	AGCTGCTGGGGGCT	GTATACGTGTCG	AGAGAGCTCAGTA	2688
Db	2676	AGGAAAG	GTGAGCTGCCCC	CGACTGGAAGGGGGCT	GTATATGTCTGTAG	AGAGCTTATGTA	2735
Qy	2629	AGACCTGTG	CCACACTGTGAGTGTCC	GACATTTGAT	TGGGACAGGGG	AGTCCACAGT	2688
Db	2736	AGACTGTG	TTTCACTGTGAGTGTCC	GGGTCACTTGTAT	TGGGACAGGGG	AGTCCACAGT	2795
Qy	2689	TCGAAAG	CCGCTCTCTCGTGA	TATCCGCTTC	GTAAAGAACTCTT	GTGAACACAG	2748
Db	2796	CCGAAAG	CCGCTCTCTCGTGA	TATCCGCTTC	GTAAAGAACTCTT	GTGAACACAG	2855
Qy	2749	CCGGAAG	AGCCATGCGACCTGT	TGAAGAGTCACTG	AAAAAGATCATCTT	GCAGAAAGAA	2808
Db	2856	AGGAGTGG	CCACACCGCTGT	TGAAGAGTCACTG	AAAAAGATCATCTT	GCAGAAAGAA	2915
Qy	2809	CTCCGTG	GTCACCAGACGTTAC	CAAGAGGGTAA	AGATGCCATG	ACGACATGTGTCA	2868
Db	2916	CTCTGCA	CTCGCCAGACTTTAC	CAAGGGGGCA	AGATGCCATG	ACGACATGTGTCA	2975
Qy	2869	GTGAAG	TGATTTTCACTC	AGAGCTGT	GAACGCCACAG	AGACTCTCAGGTT	2928
Db	2976	GTGAAG	TGATTTTCACTC	AGAGCTGT	GAACGCCACAG	AGACTCTCAGGTT	3035
Qy	2929	CTGAAG	AGATTCCG	AGATTCGGGGG	CCGAAGAC	CAACAGATGTGTCC	2988
Db	3036	CTGAAG	AGATTCCG	AGATTCGGGGG	CCGAAGAC	CAACAGATGTGTCC	3095
Qy	2989	CCGAG	TGATCCTCC	CAAGCACACAG	AGAAAGCCAC	CCAGTTCAG	3048
Db	3096	CCGAG	TGATCCTCC	CAAGCACACAG	AGAAAGCCAC	CCAGTTCAG	3155
Qy	3049	GTGCTGT	CTAGATACAG	AAGAGAGCG	CCAGACCG	ATCTTCC	3108
Db	3156	GTGCTGT	CTAGATACAG	AAGAGAGCG	CCAGACCG	ATCTTCC	3215
Qy	3109	CAGAC	AAAGTGAAGAG	AGTCCAG	GTGCTGTCA	ACCGAGCTGT	3168
Db	3216	CAGAC	AAAGTGAAGAG	AGTCCAG	GTGCTGTCA	ACCGAGCTGT	3275
Qy	3169	CAAA	AGCATGGAAG	AGTGTGAG	GGGTGAAG	AGAGATCCG	3228
Db	3276	CGAA	AGCATGGAAG	AGTGTGAG	GGGTGAAG	AGAGATCCG	3335
Qy	3229	AAGAG	AGAGACGTTAT	GTCCGA	AGACCGGTG	CAAGAGTGTG	3288
Db	3336	AAGAG	AGAGATGTGTG	CCGAAGAG	ACCGGTG	CAAGAGTGTG	3395
Qy	3289	AGGCT	CTGAGACT	GTGACGCT	ACTCCAG	AGCTTTG	3348
Db	3396	AGGCT	CTGAGACT	GTGACGCT	ACTCCAG	AGCTTTG	3455
Qy	3349	ATGTAA	CAATGTG	CGCAGT	GTCCAG	GTATTA	3408
Db	3456	ATGTAA	CAATGTG	CGCAGT	GTCCAG	GTATTA	3515

Qy	3409	TGSCCCCTGAGTCACTCGA	AACTTGCACAGACAGTGA	CAAAATCGAAGCACTCCCTTAG	3468	
Db	3516				3575	
Qy	3469	CAGATTCAGACACTCGCAGAT	TGGGACACAGCAAGATGAA	ACCAATTCAGACCGACGACAGTA	3528	
Db	3576	CAGATTCAGACACTCCAA	ACGGGACACAGACGAGACCGT	TTCAGACCGCAGGACAGTA	3635	
Qy	3529	AAGCCACTCAGCTGTCTCAGG	CAGTACAGGTCACAGAGAGAGG	CGGCTACTGCTCAGA	3588	
Db	3636	ATGCCATTGCGCGCTT	CAGCAGTACAGGTCACAT	GAAGAGGAGGACGCTGCTGCTCAGA	3695	
Qy	3589	AAGAGGAGCCTTCGACACTAC	CTAATAATGTTCCAGCCCGAC	GAGAAACATCGGAGAGAAC	3648	
Db	3696	CGGAGGGGCCCTTCAAC	CNCATC---TAGTTTTCCAGCC	CAGGAGAACACAGAGGAAAAAC	3752	
Qy	3649	CAGGAAGAGATGTTCTTTGA	ACCTPACACAGCAAGAGCTT	ACTGCTGCAGCCGTGCCCGTTC	3708	
Db	3753	CAGGAAGGGATGTTCTAGAA	ACCACCA---CAAGCGCTG	CTGCGGGCAGTGCCTATTTC	3809	
Qy	3709	TGGCAAAAGACTGAGGTGGGT	CAGAGGGTGAAGTTGACTGGT	TGGATCGAGAGAAAAGTCA	3768	
Db	3810	TGGCAAAAGGCTGAGGTGGGT	CAAGAGGGGTGAGCTGGCC	AGTTGATGAGAGAAAAGTCA	3869	
Qy	3769	AAG-----	-----AAGAACAGGAGGT	GTGTTGTACACTCTGACCC	CAACAGTCAAA	3813
Db	3870	AAGACGGACAGTGTGTTAA	AGAACTCGAGGTGCTGTGCA	CACTGGACCCAAACAGTCAAA	3929	
Qy	3814	AGGCTGCTGATGTACATAT	PGACAGTGAAGTATGGGAGT	TGGCGGGTGTTCAGGAAAAGG	3873	
Db	3930	AGACTGCTGACTTGACACG	TGACAGTGAAGTAATGGA	AGTGGCCAGATGTCAGGAACTG	3989	
Qy	3874	AGAGTACTGAAGTGCAGAGT	CTTTAGCCTCGAGAGGGAGAG	ATGAAACTGACGTTGAAA	3933	
Db	3990	AGAGTAAATGAAGAACA	CAGAGTATTAGCCCGGAGAA	AGAGAGATGGAACCGACGTTGAAA	4049	
Qy	3934	AGGAGAAAAGGGAGACAA	AGCCAGACCAAGTGCAGT	GTGAAGAGTGCAGCAGAGAAACAGCCG	3993	
Db	4050	AGAGGGAACAAGAGAC	CAAGACCAAGCCAGTGAAG	AACATGACGAGGAAACAGCTG	4109	
Qy	3994	CTCCTGAGCATGAAGGAA	ACCTACGGGAAGCCAGTCT	CTGACACTTGACATGCCAGCTCAG	4053	
Db	4110	CTCCTGAGCATGAAGGAA	CCACCCCTAAGCCAGTCT	CTGACAGCTGACACTGCTCACTCAG	4169	
Qy	4054	AGAGGGGAGGCACTGGG	NAGCCTTGGAGGAAGCCCTT	CTCTCCAGACCAAGACAAAG	4113	
Db	4170	AGAGGGAAAAGGCAC	TGGGCAGCTTTGAAGGAAG	CCCTTCTCTCCAGACCAAGACAAAG	4229	
Qy	4114	CAGGTTGCATAGAGGTT	CAAGTTCAAGCCCTCGACA	CAACAGTCACTCAAAACAGCAGAG	4173	
Db	4230	CAGATTGCATAGAGGTT	CAAGTTCAAGCTCAGAC	CACACCAGTCACTCAAAACACCGAG	4289	
Qy	4174	CTGTGGAAAAGGTCATAG	AAAACGGTTGTGATTTTCAG	AGACAGGTTGAAAGTCCAGAGTGTG	4233	
Db	4290	CTGTGAAAAGGTCGAAGA	AAACTGTGGCAACTTCAG	AGATGGAATGGATGGAGTGTG	4349	
Qy	4234	TAGGTGCACACTTATTAC	CAGCTGAGAGTCTCTCTG	CAACGGGTGGCCACTGGACTCTTC	4293	
Db	4350	CAGGTGGC	CAATCATTACCAAGCTG	AGAACTCTCCGAAAACGGGTGGCTCA	CGGGACTCTCTC	4409
Qy	4294	AGCATGCAGAGGACACGG	TACCCCTGGGGCCTGAGT	CTCTCAGGACAGAAATCCAT	TCCCAATCA	4353
Db	4410	AGCATGGAGAGGACACG	TGCCGCCCAAGGGGCTGAGT	CTTCAGGCAGAGTCCAT	TCCCAATAA	4469
Qy	4354	TAGTAACCTCTGCTCTG	AAAGCACCCCTATCCTCTG	ACTCAACAGGAGAAAATAAG	CGCAT	4413
Db	4470	TAGTAACCTCTGCTCTG	AAAGCATCCTTACA	TTCTGACCTTCAAAAGAGAA	GAGTGAAGCAT	4529
Qy	4414	CCAGAGAGAGCCATCAG	AGGAAGGACCAAGCCAGAT	GTCTGGTCTGATGCTGACGGCA		4473
Db	4530	CCAGAAAACAGAGATCAG	ATGAAGATAACAGACCCAG	ATGCTGGTCTGATGCTCGCGCA		4589

[illegible]

```

Db 5561 GATATATTTTCTTCTAGGTTTACACATGCTTTTGTATTTATACATATATATGSGT 5620
Qy 5593 GTGACGACCATGCGACGCTTTGAAAGCTCCAGGCTCACTGTAACCTGACGACAAACAG 5652
Db 5621 GTGACGACCATGCGACGCTTTGAAAGCTCCAGGCTCACTGTAACCTGACGACAAACAG 5680
Qy 5653 ----ATTAATCTCTGCGACAGACAGCAAGTCTTTTAAAGTTTACTGATGCTTAAT 5708
Db 5681 AGAAATATACATCTCTGCGACAGAGATACAGATC-TTTTAAAGTTTACTGATGCTTAAC 5739
Qy 5709 CTGCGGCTCTGAGTCCCTGTAAGAGGTTGTTTCTTATGACAGAGGCTCAGAAATA 5768
Db 5740 CTGCGGCTCTGAGTCCCTGTAAGAGGTTGTTTCTTATGACAGAGGCTCAGAAATA 5799
Qy 5769 AAAACCCATTTTGAAGATCCAGAGATGCCAATATTAACATGATTTTCCCTT 5828
Db 5800 AAAATCCATTTTGAAGATCCAGAGATGCCAATATTAACATGATTTTCCCTT 5859
Qy 5829 TTGCTAATCCAGTCCAGGTTGGAAGAGTCTCTCTGTGTGATTAAGCCCTGCTCT 5888
Db 5860 TTTCTAATCCAGTCCAGGTTGGAAGAGTCTCTCTGTGTGATTAAGCCCTGCTCT 5919
Qy 5889 TATATATATGACCAATAGATGTCCTTACAGGCTGAGATGTTTCTTATGACAGAGGA 5948
Db 5920 TAAACATATGACCAATAGATGTCCTTACAGGCTGAGATGTTTCTTATGACAGAGGA 5978
Qy 5949 TCTGTGATGATGTTTGTGATGTTGATGTTGATGATGATGATGATGATGATGATG 6008
Db 5979 TTTGTGATGATGTTTGTGATGTTGATGTTGATGATGATGATGATGATGATGATG 6037
Qy 6009 AAGTAGTCTCTGTTCTTACAGATGTTGATGATGATGATGATGATGATGATGATG 6068
Db 6038 AAGTAGTCTCTGTTCTTACAGATGTTGATGATGATGATGATGATGATGATGATG 6097
Qy 6069 ATCTGTGCCCCCTTC-TTTAAGAACATGTTGATGATGATGATGATGATGATGATG 6127
Db 6098 ATCTGTGCCCCCTTC-TTTAAGAACATGTTGATGATGATGATGATGATGATGATG 6157
Qy 6128 TGACCACTGATTTAATAAATATTTTG 6155
Db 6158 TGACCACTGATTTAATAAATATTTTG 6185

```

```

RESULT 5
AB020886 6195 bp mRNA linear ROD 14-APR-1999
LOCUS AB020886 Mus musculus mRNA for SSeCKs, complete cds.
DEFINITION AB020886
ACCESSION AB020886.1 GI:4589831
VERSION AB020886.1
KEYWORDS SSeCKs.
SOURCE Mus musculus (strain: BALB/c, isolate: 2001 9-12 weeks male testis
tissue, lib: Clontech mouse testis library cDNA to mRNA, clone: 133.5.
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Okita, K., Kitamura, H., Iwanaga, T., Morimatsu, M. and Saito, M.
TITLE Okita, K., Kitamura, H., Iwanaga, T., Morimatsu, M. and Saito, M.
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 6195)
REFERENCE Okita, K., Kitamura, H., Saito, M. and Morimatsu, M.
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Submitted (08-DEC-1998) Keio University, School of Veterinary
Medicine, Hokkaido University, Laboratory of Biochemistry, N18, W9,
Kite-ku, Sapporo, Hokkaido 060-0818, Japan
(E-mail: v080008@vet.ec.hokudai.ac.jp, Tel: 81-11-706-5206,
Fax: 81-11-757-0703)
COMMENT Sequence updated (10-Mar-1999).
FEATURES
Location/Qualifiers
1..6195
/organism="Mus musculus"
/strain="BALB/c"

```

```

COMMENT
FEATURES
Source
1..6195
/organism="Mus musculus"
/strain="BALB/c"

```

```

/isolate="200"
/db_xref="taxon:10090"
/clone="133.5"
/sex="male"
/tissue_type="testis"
/dev_stage="9-12 weeks"
/tissue_lib="Clontech mouse testis library"
1..6195
/gene="ssecks"
140..5194
/gene="ssecks"
/feature="Sic Suppressed C Kinase Substrate"
/codon_start=1
/product="SseCKs"
/protein_id="BAA76894.1"
/db_xref="GI:4589832"
/translation="MGAGSTGEQSPSPQAEVDTPSELEISGHGPAEASGAADPAD
ADPATLIPKNGQLSANGVABQEDVHVEESDQGEVEETVGHQPAEASGAADPAD
AKEMASSTVEDITKDEQETPEIIEQIPASESVEEMQAQESQANDVGFVKVFF
VGFKFTYKDKDKNSDPTVOLITVKDQEGEASVAGDHOEGBVETVGSASSEL
KQSTEROEGLKQASSTETPLQAESQGTREBAADGEMREKPEPKLESPSPVS
NETTSTFKFTFGMGWKRKTSFKPKEDDLTSEKREKQAEKDEBEGETERPAP
AEBOEPAGETDQARLSADYEKVELPLEDQVLEALSEKCAPLATEVDEKTAHOBV
VAEVHVSIVKMTKGQGAEEVDVVEGSESLPEKLAETVEVDEAPVEELMKT
KEVCSGCDHTOLDLSPBEKMLPKPEGVISVEMLSQERIKVQSPPLKIPSSSG
LKLSGKKQKGRGGGGGDEPEYOHIOETESPESADDEKSSASSEPERIACLE
KQSEBAQOEBAEBGATSDSEKREKGTTPMASHKRYTPKRYRPRESPKRELDKY
KSAATLSTESVAGMODVAVABEBOSEPKRVDTSVEMALICYSSKARAKS
SSDEGEPRLTGDDHRAEASQKENDALPASTOEDQHGSSSPBEGEV
STWESFKRLTPPKRSKSLERAEPSGAOLASIEPSREBWSIKKFIPOKRRK
ADGKQOEAIVEDSGPEINEDPDVPAVPLSEDAVERKLEAORQENVELPOLG
AVVSEELSTIVHVSVAVIDGTRAVTSABERSPSMISAMTEPLEHAGVAVLPFG
EYERKDTTATATATTAALOTLPQGDADHDITYSVDTPSEVETALRABETTA
SGAETTDWTSVSAQSUSDPTTEATPVEVEKGMIDTEQEROTAVIOAYADTK
EDSQVPATQTLQAGAPALEKVEBEESVLAETEKADVPEBPVEATEHIAQGS
EYQVATPESLEVEPTEVDVRAATQVITKHQOIMEQAVABESSLTLDSBTNGTPIA
DSIDPNTQODEVDSDSNALIAVQKOSVTEBAAAOETGSESTPSFPAQEBHRK
KPRDVEPQOALAAAPILAKAEVGGEGAGGDEKVDGOCVKELEVPVHGPS
OKTRDLTRDSEWENACQOETNESSESTISPERKMDTVEKETEETKTAQSEHNO
ETAPVEHGHGTHPKVLTADMPHSRGRALSLBSPSLPQODRADCIIEVOYSSDTYV
TQTEAVKVEEVATSEMDSESLCACAOQLPAEKSETGAGYGLTQHGEDVTPGKPS
QAESIPITVPAPEESILHSDLOREVSASQORSDENKDPAGDPAKESAAARKILR
AEPIELLESKSNKIYQSVIOTAVDQPARTEAPETASLDQNPVMAQDSQAOQM
LDKESDLOVSPDQGLSAVQEGIAVSDSEKASSEMITTLAVSAVSEVEKL
PLOCDEKHEAADQPOHSLAKAEMADSNULTESPTNPKLTEBDALKEENKAO
TEBDLQEPKDLTES"
6195
/gene="ssecks"
/feature="20 A nucleotides"
BASE COUNT 1838 a 1451 c 1777 g 1129 t
ORIGIN
Query Match 74.0%; Score 4558; DB 10; Length 6195;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 5374; Conservative 0; Mismatches 645; Indels 189; Gaps 24;

```

```

Qy 4 GCCGCGGAGTAGAAGAGCCACTAGCCATGCGCGCGACGATTCACCGACGCGCGA 63
Db 111 GCCGCGAGAGTAGAAGAGCCACTAGCCATGCGCGCGACGATTCACCGACGCGCGA 170
Qy 64 GCCCGAGAGCGCGCGCGCGCGACGCGCGAGAGAGTGTCTGAGTGGCCATGGGC 123
Db 171 GCCCGAGAGCGCGCGCGCGCGACGCGCGAGAGAGTGTCTGAGTGGCCATGGGC 230
Qy 124 CCGCAGCTGAAGCTCTGCGAGCAGTGAAGACCCCGCGACCGGACCCCGCACCAAC 183
Db 231 CCGCAGCGGAACGCTGCGAGCAGTGAAGATCCCGCGACCGGACCCCGCACCAAC 290
Qy 184 TCCACAGAAAGATGCGACGCTGTCTTCTGTCAAGGCGGTACTGAACAAGAGATGTC 243
Db 291 TCCACAGAAAGATGCGACGCTGTCTGCGTCAATGCTGTACTGAACAAGAGATGTC 350
Qy 244 ATGTCCAGAGAGAAACCGAGAGGCGCGAGAGAAAGATGCTGTATGAGATGTTGAC 303

```


QY	2458	CCACTGTGGAAAGACTCAGGGCCAGGTGGAAATTAATAAGGAGAGACCCCTTAATGTCTCCACCG	2517
Db	2556	CCGCGGTGTAAGACTCGGGGCCGAGAGAACTAATGAGAGACACCCGACGTCCACCTG	2515
QY	2518	TCGAGCTCTGTCTGAGTATTAATGACAGTGGAGAGAGAAAGATGGA-----AGCCC	2568
Db	2616	TTGTGCTCTGTCTGAGTACGATGCGGTAGAGAGAGAAAGCTGGAAACCGAGGACCTC	2675
QY	2569	AGGGGAATACGAGCTGCCCCAGCTGCTG3GGGCTGTGTACGTGTCCGAGAGCTCACTA	2628
Db	2676	AGGAGAAAGTGGAGCTGCCCCAGCTGAAGGGGGCTGTGTATGTCTGAGAGCTTTAGTA	2735
QY	2629	AGACTTGTGTCCACACTGTGTGATGTGTGCAGTCAATTGAATGGAGCCAGGGCACTGCACAGT	2688
Db	2736	AGACTCTGGATTCAACCCGTAAAGTGTGTGCGGTCAATTGATGGAGCCAGGGCACTGCACAGT	2795
QY	2689	TCGAAGAGCGGTCTCTCTGTGTGATATTCCGCTTCGTAAACAACAACCTTTGAAACAACAG	2748
Db	2796	CCGAAGACCGTCCCTCTTCTGTGATATTCTGTCTTCCATGACAGAACTCTTGAGACACGAG	2855
QY	2749	CGGAGAAAGCCATGCCCACTGTGTAAGAGGTCACTGAAAAGACATCATTCGAAAGAAA	2808
Db	2856	AGGAGTGGCCACACACGCGCTGTGGAGAGTCACTGAAAAAGACATCATCTGCAAGACAA	2915
QY	2809	CTCTCTGTGCTCACCCAGACGTTTCCAGAGGGTAAAGATGCCATGACACATGTGTCAACA	2868
Db	2916	CTCCCTGCACTCGCCCAACCTTTACAGAGGGGGCCAAATGCTCCATGACATAGTCAACA	2975
QY	2869	GTTAAAGTGAATTTCACTCCAGAAAGCTGTGACAGCCACAGACCTTCAGAGGCTCTCCGTA	2928
Db	2976	GTTGAAGTGAATTTTCACTCCAGAAAGCTGTGACAGCCAGCAAAACACAGAGGCGCTCCGCG	3035
QY	2929	CTGAAGAAATTACCGAAGCATTCG3GGGCCGAAAGACAACAGACATGTGTCTCCGACATT	2988
Db	3036	CTGAAGAACTTTACCGAAGCATTCAGGGGCGAAGAAAGACACAGCATATGTGTCTGACATT	3095
QY	2989	CCCACTGATCTGACTCTCCCAAGACACACAGAGGAACTCCCACTTACAGAGGTAGAGA	3048
Db	3096	CCCACTGTCTCGACTCTCCCGAGACACACAGAGGAACTCCCACTTACAGAGGTAGAGA	3155
QY	3049	GTTGTGTCTTGAATACAGAAAGAGAGAGGCGCAGAGCGAGGCAATCTCCAAAGCCGATG	3108
Db	3156	GTTGTGTCTTGAATACAGAAAGAGAGAGGCGCAGAGCGAGGCGCTCTCCAAAGCCGATG	3215
QY	3109	CAGACAAGTGAAGAGAGAGTCCAGGTGCTTCAACCCAGACTGTGCGAGAAACGGGGT	3168
Db	3216	CAGACAAGTGAAGAGAGACTCCAGGTGCTTCAACCCAGACTGTGCGAGAAAGCGGGC	3275
QY	3169	CAAAAAGCTTGAGAAAGTGTGAGAGGTAGAGAGAGCTTCCAAAGTGTGCTTCCGAGA	3228
Db	3276	CGAAAAGCACTGGAAGAAAGTGGAGAGGTAGAGAGAGCTTCCAGAGTGTGCTTCCAGGA	3335
QY	3229	AAAGAGAGAGCGTTATGCTCCGAAAGAGACCCGTGCAGAGAACTGTGAAGTGTGATCTTTGAC	3288
Db	3336	AAAGAGAGAGTGTGTGCTCCGAAAGAGACCCGTGCAGAGAACTGTGAAGTGTGATCTTTGAC	3395
QY	3289	AGGGCTTGAGACTGGAACAGGCTTCTCAAGAGGCTTTGAAGTCTTGAAGTCAAGCGAG	3348
Db	3396	AGGGCTTCCGAACCTGTACAGGCTTACCCCAAGAGCTTTGAAGTCTTGAAGTCAAGAGG	3455
QY	3349	ATGTGACCAATGTGCGCACGTTGCCAGGTTATCAAGCTTCAGACGCTGATGAAACAGGCG	3408
Db	3456	ATGTGACCAATGTGCGCACCAATGCGCAGGTTATCAAGCACCGAGCGTGTGAAACAGGCTG	3515
QY	3409	TGGCCCCCTGAGTCAATCCGAAACTTTGACAGACGTGAGCAAAATGGAGCACTTCCCTTAG	3468
Db	3516	TGGCCCCCTGAGTCAATCTGAAAACCTTTGACAGACGTGAGCAAAATGGAGTACTTCCCTCG	3575
QY	3469	CAGATTTCAGACATCTGCAGATGGGACAGAGAAATGAATGAACANTGACACCCAGGACAGTA	3528
Db	3576	CAGATTTCAGACATCTCMAACGGGACAGAGAAACGAGACCGTTTGACACCCAGGACAGTA	3635

QY	3529	AAGCACTGACGTGTCAGACAGTCAACAGTCTACAGAAAGAAAGGCGGTACTGCTCAGA	3588
Db	3636	ATGCATTGCGCGGCTCAAGCAGTCAACAGTCTACGAAAGAGAGGCGTGTCTCAGA	3695
QY	3589	AAGAGAGCCTTGTGACACTACCTAAATTAATGTCCAGGCCAGAGAAACATATGGGAAAGAC	3648
Db	3686	CGAGGGGCGCTTTAACAACATC---TAGTTTTCCAGCCCAAGAAACACAGGAAAAAC	3752
QY	3649	CAGAGAGATGTTCTTGAACCTTACAACAGCAAGACTTACTGTGCAGCCGTCGCTTC	3708
Db	3753	CAGAGAGATGTTCTTAGAACCAACA---CAAGCGCTGCTGCGGGCAGTGCCTATTC	3809
QY	3709	TGGCAAAAGACTGAGGTGGGTCAAGAGGGTGAAGTTGACTGGTTGATGAGAAAAATCA	3768
Db	3810	TGGCAAAAGGCTGAGGTGGGTCAAGAGGGTGAAGCTGGCACTTTGATGAAAAAATCA	3869
QY	3769	AAG-----AAGAACAGAGAGTGTGTTGTACCTCTGACCTGACCCAAAGTCAAA	3813
Db	3870	AAGACGACAGTGTGTTAAAGAACTGAGAGTCCCTGTGCACACTGACCCAAAGTCAAA	3929
QY	3814	AGGCTGTGATGTGACATATGACAGTGAAGTATGGAGTGGCGGGGTCAAGAAAAAG	3873
Db	3930	AGACTGTGACTTGACACGTCGACAGTAGATTAATGAAGATGGCCAGATGTCAAGAAACTG	3989
QY	3874	AGAGTACTGAAGTGCAGAGCTCTTAGCTCTGAGAGAGGAGATGGAACCTACCTTTAAA	3933
Db	3990	AGAGTAATGAAGAACAGAGATTTATGCCCGAGAAAAAGAGATGGGAACCAACCTTAAA	4049
QY	3934	AGAGAAAAAGGAGACAAAGCCAGACCAAGTGAAGTGAAGAAAGTGAAGAGAAACAGCG	3993
Db	4050	AGAGAGAAACAGAACCCAAACACAGACCAAGCCAGTGAAGAACATGACAGAAACAGCTG	4109
QY	3994	CTCCGTGACACTGAAGAACCTTACGGGAGACAGTCCCTGACACTTGACATGCCCCAGCTCAG	4053
Db	4110	CTCCGTGACACTGAAGAACCCCACTTAACCAAGTCCCTGACAGCTGACATGCTCCTCAG	4169
QY	4054	AGAGGGGAGAGCACTGGGAGACCTTTGAGAGAAAGCCCTTCTCTCCCAACCAAGCAAG	4113
Db	4170	AGAGGGGAGAGGCACCTGGGCAAGCCTTTGAGAGAAAGCCCTTCTCTCCCAACCAAGCAAG	4229
QY	4114	CAGGTGATGAGAGTCAAGTTCAAAGGCTGGACACAAAGTCACTCAACACAGAGAG	4173
Db	4230	CAGATGTGATGAGAGTTCAAAGTTCAAAGCTCAAGACACACAGTCACTCAACACAGAG	4289
QY	4174	CTGTGAGAAAGGTCAATGAAGACGGTTGTGATTTCAAGACAGAGTAAAGTCCAGAGTGTG	4233
Db	4290	CTGTGAGAAAGGTGGAAGAACTGTGGCAACTTCAGAGATGAGTGAAGTTTGGAGTGTG	4349
QY	4234	TAGGTGCACTTATTAACAAGCTGAGAGTCTCTGCAACGGGTGGCACTGGAATCTTC	4293
Db	4350	CAGGTGCCATCAATTCAGAGTGAAGAACTCTCCGAAACCGGTGGCTTACGGGACTCTTC	4409
QY	4294	AGCATGCAAGAGACAGGTACCCCTGGGGCTGAAGTCTCAGGCAATCAATCCCAATCA	4353
Db	4410	AGCATGGAAGAGAACCGTCCCAAGGGGCTGAAGTCTTAGGCAAGTCCATCCCATTA	4469
QY	4354	TAGTAACTCCGTGCTCTGAAAGCACTTACATCTGACCTTACAAGAGAAATTAAGCAT	4413
Db	4470	TAGTAACTCCGTGCTCTGAAAGCATCTTACATCTGACCTTACAAGAGAGTGAAGCAT	4529
QY	4414	CCCAGAGAGACGATCAAGAGAAAGAGACAAAGCAGATCTGTGCTTGAATGCTGACGCA	4473
Db	4530	CCCAAGAACAGATCAAGATGAAGATTAACAAGCAGATCTGTGCTTGAATGCTGACGCA	4589
QY	4474	AGGAGATGACAGCAATGGAAGAAAGTCTTCAAGGCTGAACCTGAAGTCTTGAAGAACTTGA	4533
Db	4590	AGGAGATGACAGCAAGAGAGAAATCTTCAAGGCTTAACTTAAGATCTTGAAGAACTTGA	4649
QY	4534	GTTAAGAGCAAAAGATTTGTGTGAAGCTCATTTAGACACCGCTTGAACAGTTCGACGCTA	4593
Db	4650	GTTAAGAGCAATTAAGATTTTCCAGAGTGTACTCCAGACAGCGCTGACAGTTTGCACGTA	4709
QY	4594	CAGAAACAGCCCGGAAACTATGCTTATATTTACAGACCCAGGTTCTG-CATGACAG	4652

Db 4710 CAGAAAAGCCCCGAAACCCAGCTTCTGATTTACAGAAATCAGGTTCTCTGTGATGAGG 4769
 Qy 4653 CTTGACAGCAGGGAGCCCAACAGATGCTGGACAAAAATGAAAG-----AT 4697
 Db 4770 CTGACAGCAGGGAGCACAACAGATGCTGGACAAAGATGAAAGCGACTTCAAGTCTCCCC 4829
 Qy 4698 GCCAAGATGAACACCCAGTGGCCGACGCCAGAGAGACTTGCAGTCTCTGACCGTTCTG 4757
 Db 4830 CCCAAGATGAACAACTCAGTGGCCGTAGCCCCCAGGAGGACTTGGCGTTTCTGATAGTCTG 4889
 Qy 4758 GAGCATGGGGCTCAGCCTCGGAAATGCTTGGCGGCTTGCAGTTGAAAGCGCGGTGTCA 4817
 Db 4890 AAGCATGAGCAAGGCTTCAGAAATGATCACACGCTTGCAGTTGAAAGTGGCCAGTGTCA 4949
 Qy 4818 AAGTAAGCATTTGAGAAGCTGCTCCTCAACCCAAAGATGCCAAAAGGAGCATGCTGCTGAT 4877
 Db 4950 AAGAAAGTGTGAGAAGCTGCTCCTCAGTGCAAGAT-GAAAAGGAACATGCTGCTGAC 5008
 Qy 4878 GGCCTCAGCTCCAAAGCTTAGCCAGGAGGAGCCAGTGCCTCTGAAAACCTTAACCAA 4937
 Db 5009 GGGCCCCAGACCAAGCTTAGCCAAAGCGAGGCGGATGCTCTGGAATCTTAACCAA 5068
 Qy 4938 GAATCCCCAGACACCAAGCGAACAGCTAACCGAGGAGGCGATCCCCCAAAGTTGAG 4997
 Db 5069 GAGTCCCCAGACACCAAGCGAACAGCTAACCGAGGAGGAGATGCCCTGAAAGAAAGAA 5128
 Qy 4998 GTCCAGGAAGAAGAAATGTCTCAACAGTCAGTCAAGAGAACCAAGGCGCCAGGCAAGAG 5057
 Db 5129 AT-----GAACAAGGCCAGACAGNAGAG 5152
 Qy 5058 ---GACCTGAGGAGCCAAAGGGAGACTGTCGCAAAATCTTAAAGATGTAGTTGCTCATTTG 5114
 Db 5153 GACGACCTACAGGAGCCAAAGGGAGACTGTCGCAAAATCTTAAAGATGTAGTTGCTCAT-- 5210
 Qy 5115 TACATCTGTAAGACCAAGATGTGAABACAAGTACAGACCAAGATGCTGCTGTGGGACC 5174
 Db 5211 -----TGTAACACTGAATGTGAAGTGAAGTACAGAACAAAGATGCTGTTGGGACC 5264
 Qy 5175 TTGAGACCAAGATTTAGAGCCCATGAGATCCAGAGAGCAGGCGCTCCAATGATTTCCA 5234
 Db 5265 TTGAGACCAAAATTTAGAGCCCTTGAGGTGACAGAGCAGAGCCGTCCAATGATTTCAA 5324
 Qy 5235 CCNAGTAGACACCCGACAAATTTAGAGGCTTATCGGAGCTAGAGCCAGCTTAACATTT 5294
 Db 5325 CCCTGCAGAGACCCGACAAATCTGAGGCTGATCGGAGCTAGAGCCAGCTAACATTT 5384
 Qy 5295 CCTCGTTTCAAGACTGCTTTGATTTGCCCTTGATGCCGCTCGTGTATTCTTAACATTT 5354
 Db 5385 CCTCTTTTCAAGACTGCCCTTGATTTGCCCTTGGTG----- 5421
 Qy 5355 CCTCGTTTCAAGACTGCCCTTTGATTTGCCCTTGTATGCCGTCGCTGATTTGCGATTTAA 5414
 Db 5422 -----CNATGATTTTCAAGTTAA 5440
 Qy 5415 GGTCTCGGTTCTCAACTGGAACCAAT-TCTGCCATACCTAGTTTCCACTTCTCAAACTG 5473
 Db 5441 GGTCTCGTCTTCTCACTCGAACCAATGTGCGCAATACCTAGTCCCACTTCTCAAACTG 5500
 Qy 5474 GAGCATCTCTTTATGATTTATATGATGTTTATGATGTTTATGATGCTCTCTGTAACCTATT 5533
 Db 5501 GAGCCTCTCTTTATGATTTATATGATGTTTATGATGTTTATGATGCTCTCTGTAACCTATT 5560
 Qy 5534 GTATA-TTTTTTTTCTAAGCTTTAAGCACATGCTTTTTTGTATTATGCAATATATAACGGGT 5592
 Db 5561 GTATATTTTTTTTCTAAGCTTTAAGCACATGCTTTTTTGTATTATGCAATATATAACGGGT 5620
 Qy 5593 GTGAGCCATAGCGACGCTTTGAAAAGCTCCAAAGCTCAAGCTCAAGCTGTAACCTGAGCAACAG 5652
 Db 5621 GTGAGCCATAGCGAAGCTTTGAGAAGCCCCCAAGCTCAAGCTGTAACCTGAGCAACAG 5680
 Qy 5653 ----ATAACATTTCTGCAAGAGAGACAAGTCTTTTTTAAAGTTTACTGATGCTTAGAT 5708

Db 5681 AGAAATTAACATTCTCTGGCAGGAAGATACAAGTC-TTTTTAAAGTTTACTGATGCTTAACT 5739
 Qy 5709 CTGTGGCTTTTCTAGTCTCTGAAAGTGGTTGTTTCTATGCACAGGAGCTCAGAAATA 5768
 Db 5740 CTGTGGCTCGTAGTCTCTGAAAGTGGTTGTTTCTATGCACAGTGAAGCTCAGAAATA 5799
 Qy 5769 AAAACCCCATTTTGAACATCCAGGATGCCCAATATTACCATGATTTTTTCCCCCTTT 5828
 Db 5800 AAAATCCATTTTGGAGACATTCAGAAATGTCCCAACATTTACGCAACATTTTTTTTTT 5859
 Qy 5829 TTGCTAATCCAGTCCAGTTGGAAAGAGTCTCTCTGTGTCAGATTAAGCCCTGTCTCT 5888
 Db 5860 TTCTAATCCAGTCCAGTTGGAAAGAGTCTCTCTAGTGTGAGATTAAGCCCATCTCT 5919
 Qy 5889 TAATGATATGGAACAATGAGTGTGCCCTTAAGCCCATGAGATGTTTCTTAAATGCAAGAGAA 5948
 Db 5920 TAACAATATGCACAGATGAGTGTGCC-ATGCCCATGAGCTATTTCCTTAACGCAAGAGAA 5978
 Qy 5949 TCTGTGTAGTGTCTTTTGTGATTTGATCTCTTCTATGCTGGACCGAATTCATATGAGATCG 6008
 Db 5979 TTTGTGTACTTTTTT-GGATTGTACTCTTCTAGGCTGGACCGAATTCATATGAGATG 6037
 Qy 6009 AAGTGAAGTCTGTTCTTTTACAGATGTTTTCATGATGATGATGTTTCTGCTGTGTTAT 6068
 Db 6038 AAGTGAAGTCTGTTCTTTTACAGATGTTTTCATGATGATGATGTTTCTGCTGTGTTAT 6097
 Qy 6069 ATCTGTGCCCCCTTC-TTTAAGAACAAATGTTGCAATATTGTTTCTTGGATAAAATGTCATT 6127
 Db 6098 ATCTGTGCCCCCTTC-TTTAAGAACAAATGTTTACATTAATGTTTCTTGGATAAAATGTCATT 6157
 Qy 6128 TGACAACTGATTTTAAATTAACATATTG 6155
 Db 6158 TGACAACTGATTTTCAATAAAATATTG 6185
 RESULT 6
 AC115647 178954 bp DNA linear HTG 13-JUL-2002
 LOCUS Rattus norvegicus clone CH230-171G16, *** SEQUENCING IN PROGRESS
 DEFINITION ** 53 unordered pieces.
 ACCESSION AC115647
 VERSION AC115647.3 GI:21737529
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 178954)
 AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
 Barbara, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buahy, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Fallis, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
 Homs, F., Howard, S., Huber, J., Huly, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louieged, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,

Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Monabadi, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokweto, S., Ogun, M., Okunolu, G.,
Ogunyemi, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y.,
Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoshitani, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Syatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Taneer, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umanai, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Unpublished
Direct Submission
2 (bases 1 to 178954)
Worley, K.C.
Direct Submission
Submitted (22-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178954)
Worley, K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:21239989.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GPHP
Center clone name: CH230-171G16
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 142023 bases at least Q40
Consensus quality: 146696 bases at least Q30
Consensus quality: 150215 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1459: contig of 1459 bp in length
* 1559: gap of unknown length
* 1560 2931: contig of 1372 bp in length
* 2932 3031: gap of unknown length
* 3032 4420: contig of 1389 bp in length
* 4421 4520: gap of unknown length
* 4521 5800: contig of 1280 bp in length
* 5801 5900: gap of unknown length
* 5901 7559: contig of 1659 bp in length
* 7560 7659: gap of unknown length
* 7660 8887: contig of 1228 bp in length
* 8888 10297: contig of 1310 bp in length
* 10298 10397: gap of unknown length
* 10398 12090: contig of 1693 bp in length
* 12091 12190: gap of unknown length

12191 13522: contig of 1332 bp in length
* 13523 13622: gap of unknown length
* 13623 15629: contig of 2007 bp in length
* 15630 15729: gap of unknown length
* 15730 17540: contig of 1811 bp in length
* 17541 17640: gap of unknown length
* 17641 19891: contig of 2179 bp in length
* 19892 19919: gap of unknown length
* 19920 21261: contig of 1342 bp in length
* 21262 21361: gap of unknown length
* 21362 23785: contig of 2424 bp in length
* 23786 23865: gap of unknown length
* 23866 25662: contig of 1677 bp in length
* 25663 25662: gap of unknown length
* 25663 27442: contig of 1780 bp in length
* 27443 27542: gap of unknown length
* 27543 30415: contig of 2873 bp in length
* 30416 30515: gap of unknown length
* 30516 31784: contig of 1269 bp in length
* 31785 31884: gap of unknown length
* 31885 34970: contig of 3086 bp in length
* 34971 35070: gap of unknown length
* 35071 36233: contig of 1163 bp in length
* 36234 36333: gap of unknown length
* 36334 39668: contig of 3335 bp in length
* 39669 39768: gap of unknown length
* 39769 42866: contig of 3098 bp in length
* 42867 42966: gap of unknown length
* 42967 46088: contig of 3122 bp in length
* 46089 46188: gap of unknown length
* 46189 49027: contig of 2839 bp in length
* 49028 49127: gap of unknown length
* 49128 51579: contig of 2452 bp in length
* 51580 51679: gap of unknown length
* 51680 55192: contig of 3513 bp in length
* 55193 55292: gap of unknown length
* 55293 58204: contig of 2912 bp in length
* 58205 58304: gap of unknown length
* 58305 60939: contig of 2635 bp in length
* 60940 61039: gap of unknown length
* 61040 64184: contig of 3145 bp in length
* 64185 64284: gap of unknown length
* 64285 67776: contig of 3492 bp in length
* 67777 67876: gap of unknown length
* 67877 72019: contig of 4143 bp in length
* 72020 72119: gap of unknown length
* 72120 75147: contig of 3028 bp in length
* 75148 75247: gap of unknown length
* 75248 78405: contig of 3158 bp in length
* 78406 78505: gap of unknown length
* 78506 82238: contig of 3733 bp in length
* 82239 82338: gap of unknown length
* 82339 85153: contig of 2815 bp in length
* 85154 85253: gap of unknown length
* 85254 87767: contig of 2514 bp in length
* 87768 87867: gap of unknown length
* 87868 90643: contig of 2776 bp in length
* 90644 90743: gap of unknown length
* 90744 94013: contig of 3270 bp in length
* 94014 94113: gap of unknown length
* 94114 98293: contig of 4180 bp in length
* 98294 98393: gap of unknown length
* 98394 101682: contig of 3289 bp in length
* 101683 101782: gap of unknown length
* 101783 104237: contig of 2455 bp in length
* 104238 104337: gap of unknown length
* 104338 108706: contig of 4369 bp in length
* 108707 108806: gap of unknown length
* 108807 113208: contig of 4402 bp in length
* 113209 113308: gap of unknown length
* 113309 117238: contig of 3930 bp in length
* 117239 117338: gap of unknown length
* 117339 122577: contig of 5239 bp in length

FEATURES									
Source									
1. .178954									
Query Match									
Best Local Similarity 72.1%; Score 4441.8; DB 2; Length 178954;									
Matches 4517; Conservative 0; Mismatches 34; Indels 6; Gaps 5;									
QY	297	GTTCGACGACGAGT	CAGAAAGATGT	CAGAGAAAGACCGAGTTGAAGAAATGGCGGCC	356				
Db	174384	GTTCGACGACGAGT	CAGAAAGATGT	CAGAGAAAGACCGAGTTGAAGAAATGGCGGCC	174443				
QY	357	AATCCACAGCTGT	TGAAGATATCA	CAAAAGGATGGCGAGGAGACATCAGAAATAATT	416				
Db	174444	AATCCACAGCTGT	TGAAGATATCA	CAAAAGGATGGCGAGGAGACATCAGAAATAATT	174503				
QY	417	GAACAGATCCCTG	TCAGAAAACAT	TGTGAAGAAATGTCACAGCTGTGTGATCCAG	476				
Db	174504	GAACAGATCCCTG	TCAGAAAACAT	TGTGAAGAAATGTCACAGCTGTGTGATCCAG	174563				
QY	477	GCTAATGATCTGG	CTTCAAGAAAGTAT	TTAAATTTCTGGTTTAAATTCACGGTGAAG	536				
Db	174564	GCTAATGATCTGG	CTTCAAGAAAGTAT	TTAAATTTCTGGTTTAAATTCACGGTGAAG	174623				
QY	537	AAGGATAAAATGA	AAAGTCAGATAC	TGTCCTCAACTACTGTCAAGAAAGGATGAAGGC	596				
Db	174624	AAGGATAAAATGA	AAAGTCAGATAC	TGTCCTCAACTACTGTCAAGAAAGGATGAAGGC	174683				
QY	597	GAAGGGCAGAAC	CTCTGTGAGCTGG	AGACACACAGAGCCAGTGTGGAGACTGCC	656				
Db	174684	GAAGGGCAGAAC	CTCTGTGAGCTGG	AGACACACAGAGCCAGTGTGGAGACTGCC	174743				
QY	657	GTCCGAGAGTCAG	CATCCAAAGAAAGT	GAGCTGAAGCAATCCACAGAGAACGAGAGGC	716				
Db	174744	GTCCGAGAGTCAG	CATCCAAAGAAAGT	GAGCTGAAGCAATCCACAGAGAACGAGAGGC	174803				
QY	717	ACCTTGAAGCAAG	ACACAGACACAGAA	ATCCCCCTTCAAGCCGAATCTGATCAAGCG	776				
Db	174804	ACCTTGAAGCAAG	ACACAGACACAGAA	ATCCCCCTTCAAGCCGAATCTGATCAAGCG	174863				
QY	777	GCTGAGAGAAAG	CCAAAGATGAAG	GAAGAAAGAAACAGAGAAAGCCCAACAAATCC	836				
Db	174864	GCTGAGAGAAAG	CCAAAGATGAAG	GAAGAAAGAAACAGAGAAAGCCCAACAAATCC	174923				
QY	837	CCAGAAATCCCC	CGAGAGCCAGTCA	CAGTGAACATCTTCTTCAAGAGTTCTTC	896				
Db	174924	CCAGAAATCCCC	CGAGAGCCAGTCA	CAGTGAACATCTTCTTCAAGAGTTCTTC	174983				
QY	897	ACTCAGGTTGGG	CCGCTGGCGCAAG	ACACAGCTTCAAGAAATCAAAAGAGGATGAT	956				
Db	174984	ACTCAGGTTGGG	CCGCTGGCGCAAG	ACACAGCTTCAAGAAATCAAAAGAGGATGAT	175043				
QY	957	CTCGAAATCTCC	GAGAGAAAGGAC	AAAGTAGACGAGGAAGAAAG	1016				
Db	175044	CTCGAAATCTCC	GAGAGAAAGGAC	AAAGTAGACGAGGAAGAAAG	175103				
QY	1017	GAAGAACACAG	ACCGCTCGAGGAG	CAGGACCGCAGACACAGACCGAGGCAGG	1076				
Db	176243	GAAGAACACAG	ACCGCTCGAGGAG	CAGGACCGCAGACACAGACCGAGGCAGG	176243				
Db	175104	GAAGAACACAG	ACCGCTCGAGGAG	CAGGACCGCAGACACAGACCGAGGCAGG	175163				
QY	1077	TTCTCAGACACT	TACAGAAAGTGG	AGCTGCCCTTTTGAAGACAGAGTTGGTGCCTGGAG	1136				
Db	175164	TTCTCAGACACT	TACAGAAAGTGG	AGCTGCCCTTTTGAAGACAGAGTTGGTGCCTGGAG	175223				
QY	1137	GCATCTCAGAG	AGAAAGTGTCTCT	CTTTGGCAACCGAAGTGTTCATGAGAAAGATGGAA	1196				
Db	175224	GCATCTCAGAG	AGAAAGTGTCTCT	CTTTGGCAACCGAAGTGTTCATGAGAAAGATGGAA	175283				
QY	1197	GCCCAACCAAG	AGTTGTCAGAG	GTCCAGTCAGACCGTGGAGAAAGACAGAGGAGGAG	1256				
Db	175284	GCCCAACCAAG	AGTTGTCAGAG	GTCCAGTCAGACCGTGGAGAAAGACAGAGGAGGAG	175343				
QY	1257	CAGGAGGAGG	AGGAGAGGCT	GAAAGGGGCGTGGTGAAGAAACAGGAGAAATCCTTG	1316				
Db	175344	CAGGAGGAGG	AGGAGAGGCT	GAAAGGGGCGTGGTGAAGAAACAGGAGAAATCCTTG	175403				
QY	1317	CCCCCTGAGAA	ACTGGCTGAGCC	CCAGGAGGTCCCCCAGGAAGCTGAGCCTGTGAGGAG	1376				
Db	175404	CCCCCTGAGAA	ACTGGCTGAGCC	CCAGGAGGTCCCCCAGGAAGCTGAGCCTGTGAGGAG	175463				
QY	1377	CTGATGAAG	AGACAGAGATGT	GTCTCTGGAGGAGACACACTCAACTGACAGACCTA	1436				
Db	175464	CTGATGAAG	AGACAGAGATGT	GTCTCTGGAGGAGACACACTCAACTGACAGACCTA	175523				
QY	1437	AGTCTGAAG	AGAGACGCTG	CGCCCAACACCCAGAAAGGCAATTCAGTGAAGTGGAGATG	1496				
Db	175524	AGTCTGAAG	AGAGACGCTG	CGCCCAACACCCAGAAAGGCAATTCAGTGAAGTGGAGATG	175583				
QY	1497	CTGCTCTCAG	AAAGAAATCA	AGGTACAGGAAAGTCCCTTGAAGAAACCTCTTCAGTAGC	1556				
Db	175584	CTGCTCTCAG	AAAGAAATCA	AGGTACAGGAAAGTCCCTTGAAGAAACCTCTTCAGTAGC	175643				
QY	1557	TCAGGCTTAA	AGAAAGCTGT	CTGGAGAGACGAGAGGGAACGAGAGGTGGGGAGAC	1616				
Db	175644	TCAGGCTTAA	AGAAAGCTGT	CTGGAGAGACGAGAGGGAACGAGAGGTGGGGAGAC	175703				
QY	1617	GAAGAGCTGG	AGAAATACCA	ACATTCACACCGAATCCCCAGAGAGTGTGATGAGCAG	1676				
Db	175704	GAAGAGCTGG	AGAAATACCA	ACATTCACACCGAATCCCCAGAGAGTGTGATGAGCAG	175763				
QY	1677	AAGGAGAGAG	CTCTGCTGCT	CCCCCGAGGAGCCTGAGAGACACAGCTGTCTGGAGAAA	1736				
Db	175764	AAGGAGAGAG	CTCTGCTGCT	CCCCCGAGGAGCCTGAGAGACACAGCTGTCTGGAGAAA	175823				
QY	1737	GGCCCGCTGA	AGACACCC	CAGGATGGGAGCTGAGAGAGGAATCTTCCGATGGAGAG	1796				
Db	175824	GGCCCGCTGA	AGACACCC	CAGGATGGGAGCTGAGAGAGGAATCTTCCGATGGAGAG	175883				
QY	1797	AAGAGAGAGA	AGGGGATCACT	CTCCCTGGGATCCTTCAAAAAGATGGTGACACCCCAAGAAA	1856				
Db	175884	AAGAGAGAGA	AGGGGATCACT	CTCCCTGGGATCCTTCAAAAAGATGGTGACACCCCAAGAAA	175943				
QY	1857	CGGGTCCGA	AGACCTTCT	GAGAGTGAAGGAGAAAGCTGGAGAAAGGTCAAGAGCGCC	1916				
Db	175944	CGGGTCCGA	AGACCTTCT	GAGAGTGAAGGAGAAAGCTGGAGAAAGGTCAAGAGCGCC	176003				
QY	1917	ACCTTCTCTC	CACTGATAG	CAAGTGTGAGAAATGCAAGATGAAATCTTCTTGGT	1976				
Db	176004	ACCTTCTCTC	CACTGATAG	CAAGTGTGAGAAATGCAAGATGAAATCTTCTTGGT	176063				
QY	1977	GAGGAAACAAA	AGCCAGAGGAA	CCAAAGCGTAGGGTGATCTTTCAGTGTCTTGGGAAGCA	2036				
Db	176064	GAGGAAACAAA	AGCCAGAGGAA	CCAAAGCGTAGGGTGATCTTTCAGTGTCTTGGGAAGCA	176123				
QY	2037	CTGATTTGTG	TGTCGGATCAT	CCAAAGAGAGACCAAGGAAGGCATCTCTTTCAGATGATGA	2096				
Db	176124	CTGATTTGTG	TGTCGGATCAT	CCAAAGAGAGACCAAGGAAGGCATCTCTTTCAGATGATGA	176183				
QY	2097	GGAGGCCAAG	AGACACTTGG	GAGGGGACAGTCAAGAGGAGGAGCCAGCAAGACAA	2156				
Db	176184	GGAGGCCAAG	AGACACTTGG	GAGGGGACAGTCAAGAGGAGGAGCCAGCAAGACAA	176243				

QY 2157 GAAGCCGGAACAGACGCTGTCTCTGCGACGACCCAGACGACGACCAAGCAAGAGT 2216
Db 176244 GAAGCCGGAACAGACGCTGTCTCTGCGACGACCCAGACGACGACCAAGCAAGAGT 176303
QY 2217 TCCCTACCCGACGACGCGGGAAGCCCTTCCGAAGGGAAGGTCCTCCACTTGGAGTCA 2276
Db 176304 TCCCTACCCGACGACGCGGGAAGCCCTTCCGAAGGGAAGGTCCTCCACTTGGAGTCA 176363
QY 2277 TTTAAAGATTAGTCACTCCAGAAAGAAAAATCCAGTCAAACTGGAAAGAAAGCCGAA 2336
Db 176364 TTTAAAGATTAGTCACTCCAGAAAGAAAAATCCAGTCAAACTGGAAAGAAAGCCGAA 176423
QY 2337 GACTCTAGTGAAGCACTGTCTCACTGAGATCGAACCGAGTGAAGAAATCTTGGT 2396
Db 176424 GACTCTAGTGAAGCACTGTCTCACTGAGATCGAACCGAGTGAAGAAATCTTGGT 176483
QY 2397 TCCATTAAAGAAATTCATCCCGGACGCGGGAAGAAAGGCGACGCGGAGCAAGAA 2456
Db 176484 TCCATTAAAGAAATTCATCCCGGACGCGGGAAGAAAGGCGACGCGGAGCAAGAA 176543
QY 2457 GCCACTGTGAAGACTCAGGCGCACTGAGATTAATGAGACGACCTTATGTCCAGCC 2516
Db 176544 GCCACTGTGAAGACTCAGGCGCACTGAGATTAATGAGACGACCTTATGTCCAGCC 176603
QY 2517 GTCTGCTCTGTCTGAGTATATCAGTGAAGAGGAGAGATGGAAGCCAGGGGAA 2576
Db 176604 GTCTGCTCTGTCTGAGTATATCAGTGAAGAGGAGAGATGGAAGCCAGGGGAA 176663
QY 2577 ACGAGCTGCCCCAGCTCTGGGGGCTGTGTAAGTCCGAGAGCTCAGTAAGCTCTG 2636
Db 176664 ACGAGCTGCCCCAGCTCTGGGGGCTGTGTAAGTCCGAGAGCTCAGTAAGCTCTG 176723
QY 2637 GTCCACATGTGAAGTGCAGATCATTTGATGGACCAAGGCGAGTCAAGTGTCAAGAG 2696
Db 176724 GTCCACATGTGAAGTGCAGATCATTTGATGGACCAAGGCGAGTCAAGTGTCAAGAG 176783
QY 2697 CCGTCTCTCTGTGATATCCGCTTCCGTAACAAGAACTCTTGAACACACAGCGGAGAA 2756
Db 176784 CCGTCTCTCTGTGATATCCGCTTCCGTAACAAGAACTCTTGAACACACAGCGGAGAA 176843
QY 2757 GCCATGCCACCTGTTGAAGAGTCACTGAAAGAAACATCTTGAAGAACTCTG 2816
Db 176844 GCCATGCCACCTGTTGAAGAGTCACTGAAAGAAACATCTTGAAGAACTCTG 176903
QY 2817 CTCAACCCAGACCTTACAGAGGTTAAAGATGCCATGACGATGTCTCAAGTGAAGT 2876
Db 176904 CTCAACCCAGACCTTACAGAGGTTAAAGATGCCATGACGATGTCTCAAGTGAAGT 176963
QY 2877 GATTTCACCTCAGAAAGTGTGACAGCCACAGAGCTCTCAGAGCTCTCCGTACGAA 2936
Db 176964 GATTTCACCTCAGAAAGTGTGACAGCCACAGAGCTCTCAGAGCTCTCCGTACGAA 177023
QY 2937 GTTACCGAAGCATTCGGGGGCCGGAAGAGACCAAGACATGTGTCCGACGTTCCAGCTG 2996
Db 177024 GTTACCGAAGCATTCGGGGGCCGGAAGAGACCAAGACATGTGTCCGACGTTCCAGCTG 177083
QY 2997 ACTGACTCCCGACACACAGAGAAAGCCACCCAGTTCAAGAGGTGAAGATGTGTG 3056
Db 177084 ACTGACTCCCGACACACAGAGAAAGCCACCCAGTTCAAGAGGTGAAGATGTGTG 177143
QY 3057 CTAGATACAGAAAGAGAGCGCCAGACGACGAGCCATCTCTCAAGCCGTTGACAGCAAG 3116
Db 177144 CTAGATACAGAAAGAGAGCGCCAGACGACGAGCCATCTCTCAAGCCGTTGACAGCAAG 177203
QY 3117 GTGAAAGAGAGTCCCAAGTGTCTGCAACCCAGATCTGTCAAGAAAGCGGGTCAAAAGCA 3176
Db 177204 GTGAAAGAGAGTCCCAAGTGTCTGCAACCCAGATCTGTCAAGAAAGCGGGTCAAAAGCA 177263
QY 3177 CTGAGAGAGTGAAGAGTGAAGAGAGACTCCGAAGTGTGCTTCCGAGAAAGAGAG 3236
Db 177264 CTGAGAGAGTGAAGAGTGAAGAGAGACTCCGAAGTGTGCTTCCGAGAAAGAGAG 177323

QY 3237 GACGTTATGCCAAAGAGACCCGTGACAGAGCTGAGCTGAGCATCTTGAACAGGGCTCT 3296
Db 177324 GACGTTATGCCAAAGAGACCCGTGACAGAGCTGAGCTGAGCATCTTGAACAGGGCTCT 177383
QY 3297 GAGACTGAGACGCTACTCCAGAGAGCTTGAAGTCTTGAAGTCAAGGATGTAGAC 3356
Db 177384 GAGACTGAGACGCTACTCCAGAGAGCTTGAAGTCTTGAAGTCAAGGATGTAGAC 177443
QY 3357 CATGTGCCACGTCGCGAGTATCAAGCTCCAGAGCTGATGGAACAGGGCGTGGCCCT 3416
Db 177444 CATGTGCCACGTCGCGAGTATCAAGCTCCAGAGCTGATGGAACAGGGCGTGGCCCT 177503
QY 3417 GAGTCATCCGAAACCTTGAAGACAGTGAAGCAATGGAAGCACTCTCCCTTGAAGATTA 3476
Db 177504 GAGTCATCCGAAACCTTGAAGACAGTGAAGCAATGGAAGCACTCTCCCTTGAAGATTA 177563
QY 3477 GACACTGAGTGGGACACAGCAAGATGAAACATTTGACAGCCAGACAGTAAAGCCACT 3536
Db 177564 GACACTGAGTGGGACACAGCAAGATGAAACATTTGACAGCCAGACAGTAAAGCCACT 177623
QY 3537 GCAGCTGACAGGCTGACAGCTGACAGAAAGAGAGCGCTACTGCTCAAGAAAGAGAG 3596
Db 177624 GCAGCTGACAGGCTGACAGCTGACAGAAAGAGAGCGCTACTGCTCAAGAAAGAGAG 177683
QY 3597 CCTTGACACTACCTTAATATATGTTCCAGCCAGAGAAACATGCGGAAAGAACAGAA 3656
Db 177684 CCTTGACACTACCTTAATATATGTTCCAGCCAGAGAAACATGCGGAAAGAACAGAA 177743
QY 3657 GATGTTCTTGAACCTTACACAGCAAGAGTTATCTGTCAGCCGTCGCTTCTGCAAG 3716
Db 177744 GATGTTCTTGAACCTTACACAGCAAGAGTTATCTGTCAGCCGTCGCTTCTGCAAG 177803
QY 3717 ACTGAGTGGGTCAAGAGGTTGAGTGAATCTGTGATGGAAGAAAGTCAATATGAC 3776
Db 177804 ACTGAGTGGGTCAAGAGGTTGAGTGAATCTGTGATGGAAGAAAGTCAATATGAC 177863
QY 3777 CAGAGAGTGTGTAACACTCTGACCCCAAGTCAAGAAAGCTGTGATGATGATGAC 3836
Db 177864 CAGAGAGTGTGTAACACTCTGACCCCAAGTCAAGAAAGCTGTGATGATGATGAC 177923
QY 3837 AGTGAAGTGAAGAGTGGCGGGGTGTCAGAAAGAGAGAGTCAAGTCTT 3896
Db 177924 AGTGAAGTGAAGAGTGGCGGGGTGTCAGAAAGAGAGAGTCAAGTCTT 177983
QY 3897 AGCTTGAGAGGAGGAGATGGAACCTGACGTTGAAAGAGAAAGGAGCAAGAGCA 3956
Db 177984 AGCTTGAGAGGAGGAGATGGAACCTGACGTTGAAAGAGAAAGGAGCAAGAGCA 178043
QY 3957 GAGCAAGTGAAGAGAGTGAAGAGAAACAGCCGCTCTGAGCATGAAGAACTTAC 4016
Db 178044 GAGCAAGTGAAGAGAGTGAAGAGAAACAGCCGCTCTGAGCATGAAGAACTTAC 178103
QY 4017 GGGAGCCAGTCTTGAACCTTGAACATGCCCAGCTCAGAGAGGGGAGGCACTGGAAAGC 4076
Db 178104 GGGAGCCAGTCTTGAACCTTGAACATGCCCAGCTCAGAGAGGGGAGGCACTGGAAAGC 178163
QY 4077 CTGGAAGAGAGCTTCTCTCCCAAGCAAGAAAGAGTGTGATGAGTCTTCAAGT 4136
Db 178164 CTGGAAGAGAGCTTCTCTCCCAAGCAAGAAAGAGTGTGATGAGTCTTCAAGT 178223
QY 4137 CAAGAGCTGGAACAGAGTCACTCAACAGCAGAGAGCTGTGAAAGAGTCAAGAAAGC 4196
Db 178224 CAAGAGCTGGAACAGAGTCACTCAACAGCAGAGAGCTGTGAAAGAGTCAAGAAAGC 178283
QY 4197 GTTGTGATTTCAAGACAGGTGAAGTCCAGAGTGTGATGACACTTATTTACAGCT 4256
Db 178284 GTTGTGATTTCAAGACAGGTGAAGTCCAGAGTGTGATGACACTTATTTACAGCT 178343
QY 4257 GAGAAAGTCTCTGCAACGGGTGCCACTGTGATCTTTCAGCATGCAAGAGCAACGGTACC 4316
Db 178344 GAGAAAGTCTCTGCAACGGGTGCCACTGTGATCTTTCAGCATGCAAGAGCAACGGTACC 178403
QY 4317 CTGGGGCTGAGTCTCAGGCAAGATTCATCCATATCATATCTCTGCTCTGAAAGC 4376

Db 178404	CTGGGGCTGAGTCTCAGCGAGAATCCCATCCCAATCANTAGTAATCTCTGCTCTCGAAGC	178463
Qy 4377	ACCTTACATCTGACCTACAAGAGGAAAAAAGCGCATCCCGAGAGAGCGATCAGAGAA	4436
Db 178464	ACCTTACATCTGACCTACAAGAGGAAATAGCGCATCCCGAGAGAGCGATCAGAGAA	178523
Qy 4437	GAGGACAAGCCAGATGCTGTGCTGATGCTGACGGCAAGGAGGTACAGCAATCGAAAA	4496
Db 178524	GAGGACAAGCCAGATGCTGTGCTGATGCTGACGGCAAGGAGGTACAGCAATCGAAAA	178583
Qy 4497	GTCTCAAGGCTGACCTGAGATCTTGAACTTGAGAGTAAGAGCAACAAGATTGTGCTG	4556
Db 178584	GTCTCAAGGCTGAACTTGAGATCTTGAACTTGAGAGTAAGAGCAACAAGATTGTGCTG	178643
Qy 4557	AACGTCATTACAGACAGCCGTTGACCAAGTTCGACGCTACAGAAACAGCCCCGAAACTCAT	4616
Db 178644	AACGTCATTACAGACAGCCGTTGACCAAGTTCGACGCTACAGAAACAGCCCCGAAACTCAT	178703
Qy 4617	GCTTATGATTACAGACCCAGGTTCTCTGC-ATGCAAGGCTTGACAGCAGGAGGCCAACAG	4675
Db 178704	GCTTATGATTACAGACCCAGGTTCTCTGCATTCAGAGGTGACAGCAGGAGGCCAACAG	178763
Qy 4676	ATGCTGGACAAAATGAAGATGCCAAGATGAACACCCAGTCCCGCAGCCAGAGAGA	4735
Db 178764	ATGCTGGACAAAATGAAGATGGCAAGATGAACACCCAGTCCCGCAGCCAGAGAGA	178823
Qy 4736	CTTCAAGTCTGACCCGTTCTGGAGGATGGCTCAGCTCGGAATCTTGGCGCGCTT	4795
Db 178824	CTTCAAGTCTGACCCGTTCTGGA-GCATGGCTCAGCTCGAATCTTGGCGCGC-T	178881
Qy 4796	GCAGTTGAAACGCGCGGTGTCAAAGTAAGCATGTGAAAGTGCCTCTCAACCCAAA	4852
Db 178882	GCANNTGAAAGCGCC-GTGTCAAAGTA--GCATGAGAGGCTGCCTCTCACCCAGA	178935
RESULT 7		
AF326230	5902 bp mRNA linear ROD 16-JUL-2001	
LOCUS	Mus musculus testis-specific SSeCKs mRNA, complete cds.	
DEFINITION		
ACCESSION	AF326230	
VERSION	AF326230.1 GI:13236407	
KEYWORDS		
SOURCE	Mus musculus.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE	1 (bases 1 to 5902)	
	Camus,A., Mesbah,K., Rallu,M., Babinet,C. and Barra,J.	
	Gene trap insertion reveals two open reading frames in the mouse	
	SSeCKs gene: the form predominantly detected in the nervous system	
	is suppressed by the insertion while the other, specific of the	
	testis remains expressed	
JOURNAL	Mech. Dev. 105 (1-2), 79-91 (2001)	
MEDLINE	21322683	
PUBMED	11429284	
REFERENCE	2 (bases 1 to 5902)	
AUTHORS	Mesbah,K., Babinet,C. and Barra,J.	
TITLE	Direct Submission	
JOURNAL	Submitted (06-DEC-2000) Immunologie, Institut Pasteur, 25 rue du	
FEATURES	Docteur Roux, Paris 75015, France	
source	Location/Qualifiers	
	1..5902	
	/organism="Mus musculus"	
	/db_xref="taxon:10090"	
	/chromosome="10"	
	/tissue_type="testis"	
	160..4899	
	/codon_start=1	
	/product="testis-specific SSeCKs"	
	/protein_id="AAK16152.1"	
	/db_xref="GI:13236408"	
	/translations="MAASSTVSDITKDEQETPEITIEQIPASESVNEEAQAASQES"	
CDS		

NDVGFKKFKFVGFKFTVKKDKNKNKSDTKVQLLLTVKKDEGEAAASVYGADGHQBPGFVEI VSESASGSELQSTQSBQSTLQKAQSSTEIPLQAESGGCTTEEAADGGEENREKPT KPLESPTSVSNVTSSTSKFFFTHWAGWRKKTSPKKPKEDDLLETSEPKCEQAELKQV EEBGEKTEPAPAEQEQPARGTDOARLSADYEKVELPLEDOVGDLLEALSEKCAPLATEV FDEKTAHQEVVAEVHVSVTVMKTKQCGAAGVEGDIVVPGSGRESLPEKLAETQEVPO EABFEBELMKTEVVCVSGDQWTLQTSPEEKMLPKHPGIVSEVMLSQERIKVQGS SPULKUFSGGLKSKKQKGRGGGDEBPEGYOHIQTESPEADQBGKSSAS PEPEBIACLEKGPSPAOAEAEAGTSDGKKRREGITPMWASFKMVTPKKVRVRPS ESDGBELDKVKSATLUSTESTASQWQDEVRVAGEORSEEPKRRVDTVSWMALCV SASVKEARKASSDSGGPRITLGGDCHRAEASKDEADALPASTOQOQOAGSSSP GASKEARASSDSGGPRITLGGDCHRAEASKDEADALPASTOQOQOAGSSSP PAGSPSEGGVSTWESFKALVTPRKKSKKLEBAESDAEQULASEPERSWSWI KXPIPPRRKKRAGDQOAAVEDSGEINEDDPDPAVVPVSEYDAVREKLEAQR QENVELPOLKGAAYVSEELSTLVHTVSVAVIDTRAVTSAEERSWASMSLE HAEAGTLPVGEVTEKDIATAEAPALAAQTLPGGKAHDHDIIVTSEVDTSEAVTAET EALRAELTEASGABETDOWASVSLSDSPGDTTEATPVQVEEGMLDTEEOBRTO AVLQAVADVKEDTSQVPAOTLQORAGPKALEKVEVEEDSEVLATEKXDVPEGPVQ EABETHLAQGSSTVQATPSLSEVPEVTDVDRATTCQVIRKQMLMQAAVAPESSETIL DEENSTGLPADSDTPNGTOQDETVDSQSNAAVKSQVTEEEAAAQVAPESSTP SFPAPQREHKRCRDLVLETPALAEVQOAGCAVPILAKEVQGEAGQFQGEKVKDQCQVKE LVPEVHTGPSNCKTADLTSEVMAEVARQETESNEEQISPEKRMVTPDQKADCI KTEQASEEHQETADLTSEVMAEVARQETESNEEQISPEKRMVTPDQKADCI EVOVSSSDTPVTCQTEAAVKVETVATSEWDSLECAQAQSLPAELKSETGGTQLQ GDTVPQGPESQAESPIIVTPAPESILHSDLQREVSQKQSDNDKPDAGDGAAG KESAEARKILRAEPLELLELSKNKIVSOIOTAVDOFARTETAPATHEDLQNVQPV MQADSOCAQMLDKDSDSLQVSPQDQTLSSVAQOGLVADSDSEGMSKASEMITLAVE SALKVESVEKLPLOCKDEKHEAADGPHQSKAKAEADASGLTKESPDTNGPKLITEEG DALKEBMNKAQTEEDDLOBPBKDLTBS"		BASE COUNT1774 a1351 c1677 g1100 t	
Query Match69.9%; Score 4304; DB 10; Length 5902; Best Local Similarity86.3%; Pred. No. 0; Matches 5105; Conservative0; Mismatches 620; Indels 189; Gaps 24;			
Qy	298	TTGGACAGCAGAGTGTGAGAGAAAAGACCAGTGTGAAGAAATGTCGAGTCCCG	477
Db	110	TTGGACAGAGAGTGTGAGAGAAAAGACCAGTGTGAAGAAATGTCGAGTCCCG	169
Qy	358	ACTCCACAGCTGTGGAAGATATCACAAAGTACGGCAGGAGGAGACATCAGAAATATG	417
Db	170	GTTCACAGTGTGGAAGATATCACAAAGCAGCAGGAGGAGAAACCCGGAATAATCG	229
Qy	418	AACAGATCCCTGCTTCAGAAAAAAATGTGGAAGAAATGGTACAGCCTGCTGAGTCCCG	477
Db	230	AACAGATCCCTGCTTCAGAGAGCAATGTGGAAGAAATGGCGCAGGCTGCTGAGTCCCG	289
Qy	478	CTAATGATGTGGCTTCAAGAAAGTATTTAAATTTGTGTGTTTTTAAATTCACGGTGAAG	537
Db	290	CTAATGACGTCCGCTTCAAGAGGTATTTAAATTTGTGTGTTTTTAAATTCACGGTGAAG	349
Qy	538	AGGATAAATGAAAGTCAGATCTGTCCAACTACTCTACTCTCAGAGGAGTGAAGGCG	597
Db	350	AGGATAAATGAAAGTCAGATCCGTCAGATCCGTCAGATCCGTCAGATCCGTCAGATCCG	409
Qy	598	AAGGGCAGAAAGCTCTGTCGAGCTGGAGACCAACAGGAGGAGTGAAGTGAAGTGAAGGCG	657
Db	410	AAGGGCAGAAAGCTCTGTCGAGCTGGAGACCAACAGGAGGAGTGAAGTGAAGTGAAGGCG	466
Qy	658	TCGAGAGTCAAGTATCAAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	717
Db	467	TCGAGTCAAGTATCAAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	526
Qy	718	CCCTGAAGCAAGCAAGCAGACGACAGAAATCCCTTCAAGCCGATCTGATCAAGCGG	777
Db	527	CCCTGAAGCAAGCAGACGACAGAAATCCCTTCAAGCCGATCTGATCAAGCCGATCTGAT	586
Qy	778	CTGAGCAAGCAAGCAAGTGAAGGAGAGAAAAAACAAGAGAAAGAGCCCAACCAAGTCCC	837
Db	587	CCGAGGAGAGCAGCAAGATGGAGAGAGAAAAACCGAGAGAGAGAACCTTCCAGGCCCC	646
Qy	838	CAGAAATCCCGAGCAGCCAGTCAACAGTGAAGCAACATCTTCTTCAAGAGTCTTCA	897
Db	647	TAGAAATCCCGAGCAGCCCTGTGAGCAATGAGCAACATCTTCTTCAAGAAATCTTCA	706

QY 898 CTGACGGTGGCCGGCTGGCGCAAGAACCAAGCTTCAAGAAATCAAAAGAGTATC 957
Db 707 CTGACGGGTGGCCGGCTGGCGCAAGAACCAAGCTTCAAGAAATCAAAAGAGTATC 766
QY 958 TGGAAACTTCCGAG 1017
Db 767 TGGAAACTTCCGAG 826
QY 1018 AAAAGACAG 1074
Db 827 AAAAGACAG 886
QY 1075 GGTGTGACAG 1134
Db 887 GGTGTGACAG 946
QY 1135 AGGATTCGTGAG 1194
Db 947 AGGATTCGTGAG 1003
QY 1195 AAGCCACCAAG 1254
Db 1004 AAGCCACCAAG 1063
QY 1255 AGCAG 1314
Db 1064 GGCACACAG 1120
QY 1315 TGGCCCTCGAG 1374
Db 1121 TGGCCCTCGAG 1180
QY 1375 AGCTGATGAAG 1434
Db 1181 AGCTGATGAAG 1240
QY 1435 TAAAGTCTGAG 1494
Db 1241 TAAAGTCTGAG 1300
QY 1495 TCCGTGCTCTGAG 1554
Db 1301 TCCGTGCTCTGAG 1360
QY 1555 GCTCAGAGCTTAAAG 1611
Db 1361 GCTCAGAGCTTAAAG 1420
QY 1612 GAGACGAG 1671
Db 1421 GAGATGAAG 1480
QY 1672 AACCAAG 1731
Db 1481 AACCAAG 1540
QY 1732 AGAAAG 1791
Db 1541 AGAAAG 1600
QY 1792 GAG 1851
Db 1601 GAG 1660
QY 1852 AGAAAG 1911
Db 1661 AGAAAG 1720
QY 1912 GGGCCACCTTGTCTGAG 1971
Db 1721 GTGCCACCTTGTCTGAG 1780

QY 1972 TTGCTGAGAAACAAAG 2031
Db 1781 TTGCTGAG 1840
QY 2032 AAGCAGATTTGTGTGATCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2091
Db 1841 AAGCAGATTTGTGTGATCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900
QY 2092 ATGAAG 2151
Db 1901 ATGAAG 1960
QY 2152 ACAAG 2211
Db 1961 ACAAG 2014
QY 2212 GAAATTCCTGAG 2271
Db 2015 GAAATTCCTGAG 2074
QY 2272 AGTCATTTAAAGATTAAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2331
Db 2075 AGTCATTTAAAGATTAAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2134
QY 2332 CCGAAGACTAGTGTGAG 2391
Db 2135 CCGAAGACTAGTGTGAG 2194
QY 2392 GGGTTCCATTAAAGATTAAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2451
Db 2195 GGGTTCCATTAAAGATTAAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2254
QY 2452 AACAGCCACTGTGAG 2511
Db 2255 AACAGCCACTGTGAG 2314
QY 2512 CAGCCGTGCTGCTGTCTGAG 2563
Db 2315 CAGCCGTGCTGCTGTCTGAG 2374
QY 2564 -AGCCAG 2622
Db 2375 GAGCTCAG 2434
QY 2623 TCAGTGAAGACTGTGTGAG 2682
Db 2435 TCAGTGAAGACTGTGTGAG 2494
QY 2683 CCAGTGTGAG 2742
Db 2495 CCAGTGTGAG 2554
QY 2743 ACAAG 2802
Db 2555 ACAAG 2614
QY 2803 AAGAAACTCTGTGTGAG 2862
Db 2615 AAGAAACTCTGTGTGAG 2674
QY 2863 TCACAGTGAAGATTAAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2922
Db 2675 TCACAGTGAAGATTAAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2734
QY 2923 TCACAGTGAAGATTAAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2982
Db 2735 TCACAGTGAAGATTAAGTCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2794
QY 2983 CAGTTTCCAGTGTGAG 3042
Db 2795 CAGTTTCCAGTGTGAG 2854
QY 3043 TAGAGAGTGTGTGATTAAG 3102

QY 231 CAAGGAGATGTCATGTCCAGAGAGAAACACAGAGGG----- 268
Db 402 CAAGATGAGCTCAGCCTCCAGAGGGTGACCTAAATGGCAGAAAGAGAGCCCTGAACGGT 461
QY 269 -----GCAGGAGAAAGAGTCCGTGATGAGGATGTTGGACACGGA 308
Db 462 CAAGGAGCCCTAAACAGCCAGGAGGAGAGAGTCAATGTCACAGAGGTTGGACAGGA 521
QY 309 GAGTCAGAGATGTGAGAGAAAGACCGAGTTGAAGAAATGGCGGCAACTCCACAGCT 368
Db 522 GACTCTGAAGATGTGAGCGAAAGAGACTCCGATAAAGAGATGGCTACTAAAGTCAGCGGT 581
QY 369 GTTGAAGATATCAGAGGATGGCAGGAGAGACATCAGAAATTAATTTGAACAGATCCCT 428
Db 582 GTTCAACACATCAGAGATGGCAGGAGAGACATCCGAAATTAATTCGAAACAGATTCCT 641
QY 429 GCTTCAGAAAACATGTGGAAGAAATGGTACAGCCTCTCAGTCCCAAGGCTAATGATGTT 488
Db 642 TCTTCAGAAAGCAATTTAGAAGAGCTAACACACCCACTGATCCCAAGGCTAATGATTT 701
QY 489 GGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTTCAAGTGAAGAGGATTAATAAT 548
Db 702 GGATTTAAGAAAGTGTAAAGTTGTTGCTTTAAATTTCAATCTGTGAATAAGGATAAGACA 761
QY 549 GAAAGTTCAGATCTGTCCTCACTACTCACTGTCAAGAGGATGAAGCGAAGGGCGAGAA 608
Db 762 GAGAAGCCTGACACTGTCAGCTACTCACTGTGAAGAAATGAAGGGGAGGGAGCGACGA 821
QY 609 GCCTCTGTCGGAGCTGGAGACCAAGAGCCAGTGTGGAGACTGCCGTCCGAGAGTCA 668
Db 822 -----GGGCTGGCAACACAGAGCCCAAGCCTTGGGCT-----GGAGAAGCA 866
QY 669 GCATCCAAAGAAAGTGAAGCAATCCACAGAGAGCAAGAGGCAACCTTGAAGCAA 728
Db 867 GCATCCAAAGAAAGCGAAACCAACAATCTACAGAGAAACCCGAAGAGACCTTGAAGCGT 926
QY 729 GAACAGAGCAGCAGAGAAATCCCCCTTCAAGCCGAATCTGATCAAGCGGCTGAGGAAGAA 788
Db 927 GAGCAAGCCACGACAGAAATTTCTCCCAAGCCGAATCTGGCCAAAGCA---GTGGAGGAA 983
QY 789 GCCAAGATGAAGAGAGAAAGAAACAGAGAAAGAGCCCAAGTCCCAAGATCCCGC 848
Db 984 TGCAAGAGAGAGAGAGAGAAACAAGAAAGAAAGAACTAGCAAGTCTGCAGATCTCCG 1043
QY 849 AGAGGCCAGTCAACAGTGAGACAACATCTTCTTCAAGAAAGTTCTTCACTCAGGTTGG 908
Db 1044 ACTAGTCCCGTGACAGTGAACAGGATCAACTTCAAAAAATTTCTTCACTCAAGTTGG 1103
QY 909 GCGGCTGGCGCAAGAGACCAAGTTCAGAAATCAAGAGAGATGATCTGGAACCTGCC 968
Db 1104 GCGGCTGGCGCAAAAGACCAAGTTTCAGGAAGCCGAAGAGGATGAAGTGGAAAGCTTCA 1163
QY 969 GAGAAGAAAGAGCAGAGGAGGAGAAAGTGAACAGAGAAAGAAAGAAAGAGACAGAG 1028
Db 1164 GAGAGNAAAGGNAAGAAGCCAGAAAGAAAGTAGACACAGAAAGAGCGAAAGGCGAGAG 1223
QY 1029 CCAGCCTCGAGGA-----GCAGGAGCCGCGCAGAGAC 1061
Db 1224 GTTGCTCCGAGAACTGACCGCTCCGAGCAAGCCCAACCAAGAGCCGCGAGAAAGT 1283
QY 1062 ACAGACAGGCGCAGGTTGTGACAGACTACAGAGAGGTGAGGTGCTGCTTTGGAAAGACCAG 1121
Db 1284 GCCCAAGAGCCCGGTTATCAGGTGAATATGAGAAAGTTGAGCTGCCCTCAGAGGAGCAA 1343
QY 1122 GTTGGTACCTGGAGGATCGTCAGAGGAGAGTGTCTCTTTGGCAACGGAAGTGT 1181
Db 1344 GTCAGTGGCTCGGAGGACCTTCTGAAGAGAAACCTGCTCCGTGGCGAGCAGAAAGTGT 1403
QY 1182 GATGAGAAGATGAAGCCCAACCAAGAA---GTTGTTGCAGAGGTCCACGTGAGCACCGTG 1238
Db 1404 GATGAGAAATAGNAGTCCACCAAGAGAGAGGTTGTGCCGAGTCCACGTGAGCACCGTG 1463
QY 1239 GAGAAGACAGAGGAGGAGGAGGAGAGAGGCTGAAGGGGGCGTGGTGTAGAA 1298

Db 1464 GAGGAGAGAAACCAAGAGCAGAAA-----ACGGAGGTGGAA 1499
QY 1299 GGAACAGGAGAAATCCTTGGCCCCCTGAGAAACTGGCTGAGCCCCCAGAGGTCCCCCAGGAA 1358
Db 1500 GAAACAGCAGGGTCTGTGCCAGCTGAGAAATTTGTTGAATGATGCAGNACCTCAGGAA 1559
QY 1359 GCTGAGCCTGCTGAGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGGAGAGACCAC 1418
Db 1560 GCCGAACCTGCCAAGGAGCTGTGAAGCTCAAAAGAAACGTGTGTTTCCGAGAGGACCT 1619
QY 1419 ACTCAACTGACAGACCTAAAGTCTGAGAGAGAGACCTGCCCAACACACAGAGCATT 1478
Db 1620 ACACGGAGCTGACCTCAGTCTGATGAGAGGTGCTGTCCAACCCCCGAGAGCGTT 1679
QY 1479 GTCAAGTGAAGTGAAGATGCTGCTCTCAGGAAAGAAATCAAGGTACAGGAAAGTCCCTTG 1538
Db 1680 GTGAGTGAAGTGAAGATGCTGTCATCAGAGAGAGAAATGAAGTGCAGGAAATCCACTA 1739
QY 1539 AGAAACTCTTCAAGTGAAGCTTAAAGAGCTCTGGGAAGAGACAGAAAGGAGAA 1598
Db 1740 AGAAGCTTTTACACAGCACTGGCTTAAAGAAAGCTTTCTGGAAAGAAACAGAAAGGAA 1799
QY 1599 CGAGAGGTGGGGAGACGAGAGCTGAGAGATACCAACACATTCACACCGAATCCCCA 1658
Db 1800 AGAGGA---GGAGGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATTTCCG 1856
QY 1659 GAGAGTGTGATGAGCAGAAAGGAGAGAGCTGTGCTGCTGCCCGAGAGCTGAGGAG 1718
Db 1857 GACAGCCAGAGAGGACAAAGGCGAGAGCTGTGCTCTATCCCTCAGAGACCCGAGGAG 1916
QY 1719 ACCACGTGTCTGGAGAAAGGCGCTGGAAGCAACCCAGGATGGGAAAGCTGAGGAAGGA 1778
Db 1917 ATCAGCTGTCTGAAAAGGGCTTTAGCCGAGGTGACAGCAGGATGGGAAAGCTGAAGAAGGA 1976
QY 1779 ACTACTTCCGATGAGAGAAAGAGAGAGAGGATCACTCCTCGGAGATCTTCAAAAAG 1838
Db 1977 GCTACTTCCGATGAGAGAAAGAGAGAGGATGTCCTCTCGGAGATCATTTCAAAAAG 2036
QY 1839 ATGCTCACACCCAAAGAAACGGGTCCGAAGACCTTCTGAGAGTGCACAAGGAGAGAGCTG 1898
Db 2037 ATGCTGACGCCCAAGAAAGCGTGTAGACGGCTTCGAAAGTATTAAGAAAGATGAGCTG 2096
QY 1899 GAGAGGTCAAGAGCCCACTTTGCTCCTCACTGATAGCACAGTGTGAGAAATGCAAGAT 1958
Db 2097 GACAAGGTCAAGAGCGCTACCTTGTCTTCCACGAGAGCACAGCCTCTGAAATGCAAGAA 2156
QY 1959 GAAGTCAAAACTGTTGAGGNAACAAAGCCAGAGAACCAAGAGGTAGGTGGATACT 2018
Db 2157 GAAATGAAGGGAGCGTGGAAAGAGCCAAAGCCGAAAGCAAGGCGCAAGGTGGATACC 2216
QY 2019 TCAGTGTCTTGGGAAGCACTGATTTGTCGGATCATCCAAGAGAGAGCAAGGAAGGCA 2078
Db 2217 TCAGTATCTTGGAGCTTTAATTTGTGGGATCATCCAAGAAAGAGCAGAGAGAGG 2276
QY 2079 TCCTCTTCAGATGATGAAGAGGGCCCAAGGACACTGGGAGGGGACAGTGCACAGAGCAGAG 2138
Db 2277 TCCTCTTCTGATGAGGAAGGGGACCAAGCAATGGAGGAGACCAACAGAAAGCTGAT 2336
QY 2139 GAGGCCAGCAAGACAAAGAACCCGGAACAGAGCTGTTCTCTGCCAGACCCAGGAGCAG 2198
Db 2337 GAGGCCGGAAGAAAGACAAAGAGACGGGGAACAGCGGATCCTTGTGTTTCCCAAGAACAT 2396
QY 2199 GACCAAGCCAGGAAGTTCCTCACCCGAGCAGCGGGAGCCCTTCCGAGGGGAGGT 2258
Db 2397 GATCCAGGGCAGGGAAGTTCCTCCCGGAGCAAGCTGGAAGGCCCTTACCGAAGGGGAGGC 2456
QY 2259 GTCTCCACTTGGAGTCAATTTAAAGATTTAGTCACTCCCAAGAAAAAATCCAAAGTCAAA 2318
Db 2457 GTTTCACCTGGAGTCAATTTAAAGTTAGTCAAGCCCAAGAAAAAATCAAGTCCAAG 2516
QY 2319 CTGGAAAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTTGTCACCTAG 2366

Db 2517 CTGGAAGGAAAAAGCAAGACTCCATAGCTGGTCTGCTGTATGAACATTCCACTCCAGAC 2576
Qy 2367 ATCCAGCCGAGTAGAGAAATCTTTGGTTTTCATTAAAGAAATTCATCCCGGAGCGCG 2426
Db 2577 ACTGAACCCGGTAAAGAAATCTCGGGTCTCAATCAAGAAATTATTCCTGGACGAAGG 2636
Qy 2427 AAGAAAAAGGCAGACGGGAACCAAGAACAGAGCTGTGGAAGACTCAAGGGCCAGTGGAG 2486
Db 2637 AAGAAAAAGGCAGATGGGAAACAAAGAACAGCCCTGTTGAAGACGAGGGCCACAGGG 2696
Qy 2487 ATTAATGAGACGACCCCAATGTCCACCGCTGTGCTCTGTCTGTATATATGCACTG 2546
Db 2687 GCCAACGAGATGACTGTGATGTCCCGGCGGTGTCTCTGTCTGTGATATGATCTGTGA 2756
Qy 2547 GAGAGGGAAGATGG-----AAGCCCAAGGGAAATACGGAGCTGCCAGCTGCTG 2597
Db 2757 GAAAGGGAAGAAATGAGGACACAGAACCCCAAAAAGCGCAGACAGCCCGACAGAG 2816
Qy 2598 GGGGCTGTGTACGTTCGAGAGAGCTCAGTAAAGACTGTGTCCACACTGTGAGTGTCCGA 2657
Db 2817 GCAGCCACTGAGTGTCTCAAGAGAGCTCAGCGAGTCAAGTTCATATGATGGCAGCAGCT 2876
Qy 2658 GTCATATGATGGGACCAAGGCGAGTCAACAATGTGAAAGCGGCTCTCCCTCGGATATCC 2717
Db 2877 GTGCTGACGGGACGAGGCGAGCTACATTAATTGAAGAAAGGTCTCTCTTCTTGATATCT 2936
Qy 2718 GCTTCCTGTAACAGAACTCTTGAACACACAGCGGAGAAAGCCATGSCACCTGTGAAG 2777
Db 2937 GCTTACAGTACAGAACTCTTGAACAGATGAAGAGCGGACCTGTTAATCAGAGAG 2996
Qy 2778 GTCACTGAAAAAGACATCATTCG---AAGAAACTCTGTGTCTCAACCAAGCGTTACCA 2834
Db 2997 GTATTGGAAGGAAGTATATTCAGAAAGAAACCCCAAGGTTATCTGAACCTTGCCA 3056
Qy 2835 GAGGGTAAAGATGGCCATGACAGACATGTGACACAGTGAAGTGAATTTCACTCAGAACT 2894
Db 3057 GAGAAACAGAGAGGCGCGGGGCAACAGGTCTGTAAGTGAAGCGGAATTGACCCCGAAGCT 3116
Qy 2895 GTGACAGCCACAGAGACTCAGAGGCTCTCCGTACTGAAGAAATTACCGAAGCATCGGG 2954
Db 3117 GTGACAGGCTGCAAGAACTGCAAGGCGCATTTGGGTCCGAAGAAAGAACCGAACATCTGCT 3176
Qy 2955 GCCGAAGAGACCAACAGACATGTGTCCGAGATTTCCAGACTGATGATCCCAAGACACC 3014
Db 3177 GCTGAAGAGACCAAGAAATGTGTCAAGCAGTCTCCAGTTAACCGATCTCCAGACACC 3236
Qy 3015 ACAGAGGAAGCCACCCAGTTAGAGAGTAGAGAGTGTGTCTAGATACAGAAAGAGAG 3074
Db 3237 ACAGAGAGGAGCCATCTCGGTGCAAGAGTGAAGGTGAGTGTACTGACATTAAGAGAGCA 3296
Qy 3075 GAGCCCAAGACCCAGGCGCATCTCTCCAGCCGTTGACAGCAAGGTGAAGAGAGTCCAG 3134
Db 3297 GAGAGGCGGACCTCAAGAGGTCTCTCCAGGCAGTGGGCAAGAAAAGTGAAGAGATCCAG 3356
Qy 3135 GTGCTGTCAAC--CAGACTGTGACAGACGCGGTCAAAAAGCACTGAGAAAGTTGAG 3191
Db 3357 CTGCTCGGACCCGTTGGGCGCAGAAAGATGTCTTCAAGCTGTGACAGACAGGAGAGAA 3416
Qy 3192 GAGGTGAGAGAGACTCCCAAGTGTGCTTGGAGAAAGAGAGAGAGCTTATGCCGAAA 3251
Db 3417 AGACCAAGAGACAGGCTAAGGCTGAGGTCTGAAGAAAGAGAGAGATGATGTTGAAA 3476
Qy 3252 GGAACCCGTGACAGAGCTGAGAGCTTGTGACACAGGGCTCTGAGACTGAGACAGGCT 3311
Db 3477 GTAGATGCTCAGAGAGCAAAAACTGAGCTTTTACCAAGGAGAGGTGTGGGGGAGAGAC 3536
Qy 3312 ACTCCAGAGAGCCCTTG--AAGTTCTGAAGTCAAGGCAAGATTAAGC-----CAT 3359
Db 3537 ACCCAAGAAAGCTTTGAAAAAGCTCTCAAGTCAAGAGAGCATTAAGTCAAGTGAAGCTT 3596
Qy 3360 GTCCGCAAGCTTGCA-----GGTATCAAGCTCCAGAGAGCTGATGAAA 3401
Db 3597 GTAACCACTTGTCAAGCCGAACCTTAACTGGGGTAAAAATCACAGAGATGGTATGAAA 3656

Qy 3402 CAGGCGGTGGCCCTGAGTATCCGAAACCTTGAACAGACAGTGAACAATGGAAGCACT 3461
Db 3557 CAGGCTATCCCTCTGCTGTGGGAAACCTTACACACAGTGAAGTGAATGGAAGACCC 3716
Qy 3462 CCCTTACAGATTCAGACATGCAATGGAACACAGCAAGATGAACCATTTGACAGCCAG 3521
Db 3717 CCGGTACCGCACTTTGACGACACAGGCAACAACCAAGAAAGCGATGTGTGAATTCAT 3776
Qy 3522 GACATTAAGCCACTGCACTGTGTACAGCACTCACAGTCAACAGAAAGAGCGGCTACT 3581
Db 3777 GAGAGAGATGAGTGTGATCTGTGATCCAGTCAAGGGGCAAGAAAGCAGAGGCACTTCT 3836
Qy 3582 GCTGAGAAAGAGAGCCCTTGACACTACCTAATTAAGTTCAAGCCCGAAGAACATGGG 3641
Db 3837 GCACAGAAAGAGAGGCTCCAGACCTTCCAGTTTGTGTTCCAGAGAAACTTAAGAA 3896
Qy 3642 GAAGAACAGAGAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTACTGTGACAGCCGTG 3701
Db 3897 CAATCAAGATGGAAGACACTCTAGAGCATTAAGTAAGAGGTGTCAGTGGAAACTGTA 3956
Qy 3702 CCGGTTCTGGCAAGACTGAGTGGGTCAAGAGGTGA-----GGTTGACTGTTG 3752
Db 3957 TCCATTCTGTCAAGACTGAGGGGACTCAAGAGGCTGACAGATGTGCTGATGAAGAAAC 4016
Qy 3753 GATGAGAAAAAGTCAAAAGAAAGACAGAGGTGTTGTACACTGTGG-----ACCAAC 3806
Db 4017 AAAGACGTACATTTTTCGAAAGCTTGAAGGAGTGTATAGCACAGGCATTAACGTGAT 4076
Qy 3807 AGTCAAAAGCTGTGATGTGACATATGACAGTGAATGAGTGGAGTGGCGGCTGACAG 3866
Db 4077 CGGAAAAAGTCACTGAAGTTGCCCTTAAAGTGAAGGACAGAAAGAGTGAATGTA 4136
Qy 3867 GAAAAAGAGACTGATGAATGACAG-----TCTTAAGCTGTGAGAGGGA 3911
Db 4137 AAGGATGATGCTCTTGAACCTGACAGTCAACGTAACTCTCCATCCCGCGTGAAGAGA 4196
Qy 3912 GAGATGGAACCTGACCTTGAAGAAAGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGAA 3971
Db 4197 GAGATGATGTTCAAGTGTGAAGGAGAGAAACAGAAAGAGCCCAATGTAAATGA 4256
Qy 3972 GAAGG---TGAGCAGAGAAACAGCGCTCTGAGCATGAAGAACCTTACGGGAAGCCAGTC 4028
Db 4257 GAGAAAGCTTGAAGCAAGAAACAGCTGTATCTGTAAGAGGTCAATGAACAGCTCTC 4316
Qy 4029 CTGACATTTGACATGCCAGCTCAGAGAGGGGAAAGCACTGGGAAGCTTTGAGAGAAC 4088
Db 4317 CAGACAGTGAATGTGCCATCATAGATGGGCAAAAGAAATCAGCAGTTTGGAAAGAAC 4376
Qy 4089 CCTTCTC---TCCAGACCAAGCAAGAGCAAGTGTGACATGAAGTCAAGAGCTG 4145
Db 4377 CCTCTCTCTGCTTGAAGTCAAGAGAGGCAAGTATGACAAATTTCAAGTTCAAGGCTCT 4436
Qy 4146 GACACAAAGTCACTCAACAGCAGAAAGCTGTGAAAA-----GGTCAATGAAGCGTT 4199
Db 4437 GAGCATATTTCACTTAAACAGCGCTGACAGAGAGAGAAAGTCTTATGAGAAACTGCC 4496
Qy 4200 GTGATTTCAAGACAGGTGAAGTCCAGAGTGTGTAGGTGACACTTATTAACAGCTGAG 4259
Db 4497 AACATTTTGAAGAACAGTGAAGAGCTGAGAGCTGACAGTGAATTTAGTTCTGGAAGAG 4556
Qy 4260 AAGTCTCTGCAACGGGTGGCACTGCACTTTACAGCATGCAAGAGACACGATACCCCTG 4319
Db 4557 AAATCTCTGAAAAAAGAAAGACTTTCCTCCCTCATCCAGGGGAAAGAGCTGTGTCCCA 4616
Qy 4320 GGGCTGAGTCTCAGGAGAAATCCATCCCAATCAATGAATCTCTGCTCTGAAAGACCC 4379
Db 4617 GGGCCGAGCTGTCAAGCAAAATTCACACAGTATGATGTGTCTACTTACCAAGAAAGGC 4676
Qy 4380 CTATCTCTGACTCAAGAGAAATTAAGCGATCCACAGAGAGAGATCAAGAGAAAG 4439
Db 4677 TTAAGTTCCGACCTGGAAGGAGAGAAACCAATCATCAAGTGAAGTGAATGAAGTGC 4736

Qy	4440	GACAAGCCAGNAGTCTGTGCTGATGCTGACGGCAGGAGAGTACAGCAATCGAAAAGTC	4490
Db	4737	GATGACGAGGTGCTGTGCC-----AGSAGGTCCAAAGTGTAGTCAATTTGAGGAGGAT	4790
Qy	4500	CTCAAGGCTGAACCTGAGATCTCTGGAACTTTGAGAGTAAAGAGCAACAAGATTTGTGCTGAAC	4559
Db	4791	TTAGAGCCTGNAANTGGATTTTGGAACTTTGAGACCAAAAGCGTAAACTTTGTCCAAC	4850
Qy	4560	GTCAATTCAGACAGCCGTTGACCAAGTTGCAACGTAC---AGAAACAGCCGCCGAAACTCAT	4616
Db	4851	ATCATTCAGACAGCCGTTGACCAAGTTGTACGTACAGAAAGAACAGCACCAGCAATGTG	4910
Qy	4617	GCTTATGATTCACAGACCCAGTTTC-----CTGCATCGAGGCTTGACAG	4660
Db	4911	ACGTCTGAGTTACAGACACAAGCTCACGTGATAAAGAGCTGACAGCCAGGAGCGTGGACAG	4970
Qy	4661	CAGGAGCCCAACAGATGCTGGACAAAAATGAAGATGCCAAGATGAACAACCCAGTGCC	4720
Db	4971	GAACGGAGAAAGAGGAGGAGCACTCAGGCCTCTGCACAGGATGAACAACCAATTACT	5030
Qy	4721	GCAGCCCA-----GAGAGGACTTTGCAAGTCTTGACCCGTTCTTGGA	4759
Db	5031	TCAGCCAAAGAGGAGTCAGAGTCAACCGCAGTGGGACAAGCACATTCTTGATATTCCAAA	5090
Qy	4760	GGCATGTGGCTCAGCCTCGGNAATGCTTCGGCGCTTGCAAGTTGAAGCGCCGGTGCAA-	4818
Db	5091	GACATGAGTGAAGCCTCAGAAAAGACCATGACTGTTGAGGTAGAAAGTTCCACTGTAAAT	5150
Qy	4819	-----4818-----	4818
Db	5151	GATCAGCAGTCGNAAGAGGTGCTCTCCCATCTGAGGAAGAGGAGGTGGAGCTGGAAACA	5210
Qy	4819	-----AGTAAGCATTTGAGAAGCTGCCT	4840
Db	5211	AAGTCTGTGCAGAGATGATGTCTCATGCTTTGTAGCAGAAAGATAGAGAAGTCACTA	5270
Qy	4841	CCTCAACCCAAAAG--TCCAAAAGGAGCATGTCTGTGATGGCCCTCAGCTCCAAAGCTTA	4898
Db	5271	GTTGAACCGAAAGAAAGATGAAAAGGTGATGATGTTGATGACCTTGAAAACAGAACTCA	5330
Qy	4899	GCCAGCAGCAGAGCCAGTGCTCTGGAACCTTAAACCAAGAAATCCCCAGACACCAACGGA	4958
Db	5331	GCCCTGCTGATGATGATGCTCAGAGGCTTAAACCAAGAGTCCCCAGATACAATTTGGA	5390
Qy	4959	CCAAAGCTAACCGAGGAGGCGCATCCCCAAAGTTGAGGTTCCAGGAAGAGAAATGTCT	5018
Db	5391	CCAAAAACAAAAGAGAGGAGGATGCCAGGAAGTAGAATTTCCAGGAAGGAAGTGCAC	5450
Qy	5019	ACCAAGTTCAGTCAAGAGAAACAAGGCCAGCGCAAGAGGACCTTCAGAGGCCAAAGGGA	5078
Db	5451	AGTGAATTCAGATAAAGCGATCACACCCCAAGCACAGGAGGAGTTACAGAAAACAAGAGA	5510
Qy	5079	GACCTGCAGAACTCTTAAGATGTTAGTTG-----CTCATCTTAC	5117
Db	5511	GAATCTGCAAGTTCAGAACTTACAGAACTTTAAAACATCATGCAAGTTAACTCATGTCT	5570
Qy	5118	ATCTGTGAAGCCAGAATGTGAAAACAAGTCACAGAACAA---GATGCTGCTGTTGGGAC	5173
Db	5571	GTTTGGAAAGACCAGAATGTGAAGACAAGTAGTAGAAGAAATGAATGCTGCTGTGAGA-	5629
Qy	5174	CTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGCAGGCGCGTCCAATGATTTCC	5233
Db	5630	CTGAAGACCAAGTATTTTTCAG-ACTTTGAAGAAATTGGAGAGCAGGCACATCACTGATCTCA	5688
Qy	5234	ACCCAGTAPAGACCCCGCAATCTCTGAGGCTTCATCGGAGCTAGAGCCAGCTAAACATT	5293
Db	5689	TTTCTAGAGAGC--CCCTGACAATCTCTGAGGCTTCATCAGGAGCTAGAGCCATTTAAACATT	5747
Qy	5294	TCCTCGTTTCAAGACTGCCTTTTGATTTGCCCTCTGCATGCGCGTGTATTTCTTAACATT	5353
Db	5748	TCCTCTTTCCAAGACCAACTACATTTTCCCTTGTATTAAC-----5787	5787
Qy	5354	TCCTCGTTTCAAGACTGCCTTTTGATTTGCCCTCTGCATGCGCGTGTATTTTCGGATTTA	5413

D	b	5788	- - - - -		-----CATATAAATCTGATTTA	5805
Q	y	5414	AGGTCTCGGTTCTCAACCTGGAAC-CAATTC		TGCACATACCCTAGTTCACACTCTCAAAC	5472
D	b	5806	AGGTCTAAATCTTAACCTGGAACTGGAGTT		TGGCAATACCTAGTTCTGCTTCTGAACT	5865
Q	y	5473	GGAGATCCTCCTTATATGTTATATGTTATAT		GTTATGTTTAGTCTCCTCTGTACCTAT	5532
D	b	5866	GGAGTATCATCTTTTACATATTTATATGTTAT		GTATGTTTAAAGTAG---TCCCTCTGTATCTAT	5922
Q	y	5533	TGTATATTTTTTCTAACGGTTT-		-----AAGCACATCCTTTTTTGTTATTA	5575
D	b	5923	TGTATATTTTTTCTTAATGTTTAAAGAAATG		TGCAGGAATCTACATGCTTTTTTGATATCA	5982
Q	y	5576	TGCAATATATAACGGGTGTGCAGCCATAGCAG		CGCTTTGAAAAGCTCCAAGCCTCAAC	5635
D	b	5983	CACAGTATATGATGGGCGATGTGCCATAGTG		CGAGGCTTGGGAGCTTTAAGCCTCAGTTA	6042
Q	y	5636	TAACTCGAGCAACAGATAACATTC-		-----CTGGCAAGAAGACAAGTCTTTTT	5686
D	b	6043	TATAACCCACAAAAACAGAGCCTCCTAGATG		TAAATCTCTGTGATCAAGGTACAAATCTTT	6102
Q	y	5687	TAAAGTTTACTGATGCTTAGATCTGTGGGCT		TCTTAGTCTCTGAAAGGTGTTGTTTTCC	5746
D	b	6103	TAAATTCATNATGTTAGGTCATATTTAGTG		GTACTCTGAAATTTGGTTCACCTTTCT	6162
Q	y	5747	ATGCACGCGAGCTCAGAAATAAAACCCCATT		TTGAAACATCCAGGATGTCCCCAATATT	5806
D	b	6163	ATTACGCGAGTGTGCCAAACCTAAAAGCA		TTTTGAAACATACAGAATGTTCTATTGTC	6222
Q	y	5807	ACCATGATTTTTTCCCCCTTTTGTCTAAT		CAGTCCAGGTTGGAAAGAAGTCTCCTCTG	5866
D	b	6223	ATTGGGAAATTTTGC-----TTTCTAA		CCAGTGGAGGTTAGAAAAGAAGTTATTTCT	6275
Q	y	5867	TGTCAGA-----TTAAGCCCTGTCTT		TAATGATATGGACAATCAGT	5909
D	b	6276	GGTAGCAAAATTACTTTACATCCTTTTCT		TACTGTTATGGTGTGTTGGACCGATAGT	6335
Q	y	5910	GTGCCTAAGCCCATGAGA-----		-----TGTTTCTTAATSCAAGAGGAATCTGT	5953
D	b	6336	GTGCTTAATCCTGAGGCCAAAGTAGTGA		ATATGTTTATGTTATGAAGAAAAGAAATGT	6395
Q	y	5954	TGTAGGTTTTTTTGATTTGTACTCTTCTAT		CTGTGGACCGAATTCATATGCAGATCGAAGTG	6013
D	b	6396	TGTAAG--TTTTTGATTCTACTCTTATAT		GTGCTGCAATTCGATTCACACATGGCATGAAATA	6453
Q	y	6014	AGTCCGTCTCTTACAGATGGTATTTTGAT		TAGATACTCGAGTTGTCTGTCTGTATATCTG	6073
D	b	6454	AGTCAGGTTCTTTACAAATGGTATTTTG		ATAGTACTGATTTGTTTGCCCATATTTTG	6513
Q	y	6074	TGCCCCCTCTTTAAGAACAAATGTCATTT		ATGTCCTTTTGGATAAAATGTGATTTGACAA	6133
D	b	6514	TGCCATCTTTTAAAGAACAAATGTTGCA		CACATTCATTTGGATAAGTTGTGATTTGACGA	6573
Q	y	6134	CTGATTTAAATATAA	6147		
D	b	6574	CTGATTTAAATATAA	6587		
RESULT		9				
AX329766						
LOCUS		AX329766		6608 bp	DNA	linear
DEFINITION		Sequence 275 from Patent WO0194629.				PAT 09-JAN-

SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Young, P. E., Augustus M., Carter, K. C., Ebner R., Endress, G.,

Db 2034 ATGCTGAGCCCAAGAAGCGTGTAGAGCGCTTTCGGAAGTGATAAAGAGATGAGCTG 2093
Qy 1899 GAGAAGTCAAGAGCGCCACCTTGTCTCTCACTGATAGCACAGTGTGAGAAATCAAGAT 1958
Db 2094 GACAAGTCAAGAGCGCTTACCTTGTCTTCCACCGAGAGCACAGCTCTGAAATGCAAGAA 2153
Qy 1959 GAAGTCAAAACTGTTGGTAGGACAAAGACCAAGGACCAAGACGTTAGGTCGATACT 2018
Db 2154 GAAATGAAGAGGAGCGTGGAAAGAGCCAAAGCCGGAAGAACCAAGCGCAAGGTGGATACC 2213
Qy 2019 TCAGTGTCTTGGGAAGCACTGATTTGTGTGCGATCATCCAAAGAGAGAGCAAGAGGCA 2078
Db 2214 TCAGTATCTTGGGAAGCTTTAATTTGTGTGGATCATCCAAAGAAAGAGCAAGGAGG 2273
Qy 2079 TCCTTTCAGATGATGAAGAGGCGCAAGGACACTGGAGGGGACAGTCAAGAGCAGAG 2138
Db 2274 TCCTTTCATGATGAGGAGGGGACCAAAAGCAATGGAGGAGAGCACCAAGAAAGCTGAT 2333
Qy 2139 GAGGCCAGCAAGACAAAGAGCCGGAACAGACGCTGTTCTGCGCAGCACCCAGAGCAG 2198
Db 2334 GAGCCCGAAAGAACAAAGAGACGGGACAGACGGGATCCTTGTGTTCCCAAGAACAT 2393
Qy 2199 GACCAAGCGCAAGGAAGTTCTCAACCGAGCCAGCGGAGCCCTTCCGAAGGGGAAGGT 2258
Db 2394 GATCCAGGCGAGGAAGTTCTTCCCGAGCAAGCTGGAAGCCCTACCGAAGGGGAGGC 2453
Qy 2259 GTCTCCACTTGGGAGTCATTTAAAGATTAGTCACTCCAAAGAAAAAATCAAGTCAAAA 2318
Db 2454 GTTTCCACTTGGGAGTCATTTAAAGGTTAGTCACGCCCAAGAAAAAATCAAGTCCAAG 2513
Qy 2319 CTGGAAGAGAAAGCCGAGAC-----TCTAGTGTAGACGATGTTGTCACCTGAG 2366
Db 2514 CTGGAAGAGAAAGCGAAGACTCCATAGCTGGGTCTGGTGTAGAACATTCACCTCCAGAC 2573
Qy 2367 ATCGAAACCGATAGAGAAGATCTTGGTTTCCATTAAAGAAATTCATCCCGGACGCGG 2426
Db 2574 ACTGAAACCGGTAAAGAGATCTTGGTCTCAATCAAGAAATTTATCTCGAGAGAGG 2633
Qy 2427 AAGAAAAGGCGAGCGGGAAGCAAGAACCAAGCCACTGTGGAAGACTCAGGGCCAGTGGAG 2486
Db 2634 AAGAAAAGGCGAGATGGAAACAAAGAACAGCCCTGTGTAAGACGCGAGGSCCAACAGGG 2693
Qy 2487 ATAAATGAGACGACCTTATGTCACGCGTCTGTGCTCTGCTGATGATATATGCAAGTG 2546
Db 2694 GCCAACCAAGATGACTGATGTCGCCGCGTGTCTCTGCTGATGATGATGCTGTA 2753
Qy 2547 GAGAGGAGAGATGG-----AAGCCAGGGGAATACGGAGCTGCCACGCTCTG 2597
Db 2754 GAAAGGAGAAATGGAGGCAACAGCCCAAAAGGCGCAGAGCAGCCGCAAGAAAG 2813
Qy 2598 GGGCTGTGTACGTGTCGAGGAGCTCAGTAAGACTCTGGTCCAACACTGTGAGTGTGCGCA 2657
Db 2814 GCAGCCACTGAGGTGTCCAAGGACTCAGCGAGAGTCAAGTTTATATGATGGCAGAGCT 2873
Qy 2658 GTCAATTGATGGACAGGCGAGTCACAGTGTGGAAGCGGTCTCTTGTGATATACC 2717
Db 2874 GTCGCTGACGGGACGAGGCGAGCTACCATTTATGAAGAAAGGTCTCTCTTCTTGGATATCT 2933
Qy 2718 GCTTCGTGTACAGAACTCTTGAACACACAGCGGGAGAACCATGCCACCTGTTGAAAG 2777
Db 2934 GCTTCAGTGACAGAACTCTTGAACAAAGTAGAAGCTGAAGCCGCACTGTTAACTGAGGAG 2993
Qy 2778 GTCACTGAAAGACATCATTTGC---AGAAGAACTCTCTGTGCTCAGCCAGACGTTTACCA 2834
Db 2994 GTATTGGAAGAGAGATTAATTGCAAGAGAGAAACCCCCACGGTTACTGAACCTCTGCCA 3053
Qy 2835 GAGGTTAAAGATGCCATGACGATGTTCAACAGTGAAGTGGATTTCACTCAGAAAGCT 2894
Db 3054 GAGAACAGAGAGGCGCGGGCGACAGCGTCTGTAGTGGCGGAATTTGACCCCGAAGCT 3113
Qy 2895 GTGACGCCACAGAGACTCAGAGGCTCTCGTACTGAGAGTTTACCGAGCATCGGG 2954
Db 3114 GTGACAGCTGCAGAACTGCGAGGGCCATTGGGTTCCGAAGAGGAACCGAAGACATCTGCT 3173

Qy 2955 GCCGAAGAGACCAAGACATGGTGTCCGCAAGTTTCCAGCTGACTGACTCCCGCAGACACC 3014
Db 3174 GCTGAAGAGACCAAGAAATGGTGTGAGCAGTCTCCAGTTAAACCGACTCCCGCAGACACC 3233
Qy 3015 ACAGAGGAGCCACCCAGTTTCAGGAGGTAGAGTGGTGTCTAGATACAGAGAGAGAG 3074
Db 3234 ACAGAGGAGCCACTCCGGTGCAGGAGGTGGAAGGTGGCGTACTTGACATAGAGAGCAA 3293
Qy 3075 GAGCGCCAGACGCGAGCCATCTCTCAAGCCGTTGCGACACAAGGTGAAAGAGGATCCAG 3134
Db 3294 GAGAGCGGACTCAAGAGGTCTCTCCAGGCACTGCGAGAAAAGTGAAGAGGAATCCAG 3353
Qy 3135 GTGCTGCAACC---CAGACTGTGCGAGAAACGGGGTCAAAAGCACTGGAGAGAGTTGAG 3191
Db 3354 CTGCTGGCAACCGGTGGCCAGAAAGATGTCTTTCAGCCTGTGCGAGAGCAGAGGCGAGAA 3413
Qy 3192 GAGGTAGAGGAGGACTCCGAAGTGTGGCTTCGGAGAAAGAGAGGAGCGTTATGCCGAAA 3251
Db 3414 AGACCAAGAGAGGCTGAAGCGTCGGGTCTGAAGAAAGAGACGGATGTAGTGTGAAA 3473
Qy 3252 GGACCCGTGCGAAGCTGGAGCTGAGCATCTTGCACAGGGCTCTGAGACTGGACAGGCT 3311
Db 3474 GTAGATGCTCAGAGGCAAAAATCTGAGCTTTTACACAAGGGAAGTGGTGGGCGACACC 3533
Qy 3312 ACTCCAGAGAGCCTTG---AAGTTCTGAAAGTCAOCCGACAGATGTAGAC-----CAT 3359
Db 3534 ACCCAGAAAAGCTTGAAGAAAGCTCTCAAGTCAAGAGAGCATAGAGTCCAGTGAGCTT 3593
Qy 3360 GTGCCACGTCGCA-----GGTTATCAAGCTCCAGCAGCTGATGGAA 3401
Db 3594 GTAACCACTTGTCAAGCCGAAACCTTAGCTGGGTGTAATCAAGAGAGATGGTATGGAA 3653
Qy 3402 CAGCCGTGCGCCCTGAGTCATCCGAAACCTTGACAGACAGTGAACAAATGGAAGCACT 3461
Db 3654 CAGGCTATCCCCCTGACTCGGTGGAAACCCCTTACAGACAGTGAAGTGTGATGGAGACCC 3713
Qy 3462 CCCTTAGCAGATTCAGACACTGCAGATGGGACACAGCAAGATGAAACCATTTGACAGCCAG 3521
Db 3714 CCCGTAGCCGACTTTCAGCCACAGGCGCAACCCAGAAACAGAGATTGTGGAAATCCAT 3773
Qy 3522 GACAGTAAGCCACTCGAGCTGTGCGGAGTCAAGGTGACAGAGAGAGGCGGCTACT 3581
Db 3774 GAGGAGAAATGAGTGCATCTGTTGCCACCTTCCAGTCAAGGGGCAAGAGCAGAGGAGTTCCCT 3833
Qy 3582 GCTCAGAAAGAGGAGCTTCGACACTACCTAATAATGTTCCAGCCCAAGGAAGCAATGGG 3641
Db 3834 GCACAGAAAGAGGCTTCAGCACCTTCCAGTTTGTGTTCCAGGAAAGAACTAAAGAA 3893
Qy 3642 GAAGAAACCAAGAGAGATGTTCTTGAACCTTACAGCAAGAGCTTACTGCTGCAGCCGCTG 3701
Db 3894 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAGAGGTGTCAAGTGAAGAACTGTA 3953
Qy 3702 CCCGTTCTGCAAAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3954 TCCATTCTGTCAAAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCTGATGAGAAACCC 4013
Qy 3753 GATGGAGAAAAGTCAAAAGAAAGAGGAGGTGTTGTACACTCTG-----ACCCAAC 3806
Db 4014 AAAGACGTACCATTTTCGAAGGACTTTGAGGGTCTATAGACACAGGCATTAACAGTCAGT 4073
Qy 3807 AGTCAAAAGGCTGCTGATGTGACATATGACAGTGAAGTGTGAGGTGGCGGCTGTGAG 3866
Db 4074 CGGAAAAAGGTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAGAGTGAATGTAAA 4133
Qy 3867 GAAAAGAGAGTACTGAATGCGAGAG-----TCTTAGCCTGGAGGGGGA 3911
Db 4134 AAGGATGATGCTCTTGAACCTGCGAGAGTCAAGTCTCTCCATCCCGCTGGAGAGA 4193
Qy 3912 GAGATGGAACCTGAGTGTGAAGGAGAAAGGAGACAAAGCCAGAGCAAGTGTAGTGA 3971
Db 4194 GAGATGATGTTCAAGTTCGAAGAGGAGAAACAGAGACGAGCCCAACCCATGTGTAATGA 4253

QY 3972 GAAG---TGAGCAGAAACAGCGCTCTTGAGCATGAAGCACTACGGAAAGCACTC 4028
DB 4254 GAGAAAGCTTGAGCAGCAAGAAACAGCTGTATCCGTATCTGAAGAGTGAAGTAAGAGAGCTCTC 4313
QY 4029 CTGACATCTTGACATGCGGAGCTCAGAGAGGGGAGAGCACTGGAAAGCTTGAGAGAGC 4088
DB 4314 CAGACAGTGAATGTGCTCATATGATGGGCAAGAGAGTCAAGAGTTTGAGAGAGAGC 4373
QY 4089 CTTTCTC---TCCAGACCAAGACAAAGAGTTGCTATGAGGTTCAAGTTCAAGCTG 4145
DB 4374 CTTCTCTCCCTGCTATGCTCAAGAGAGAGCAATATGCAACAAATTCAGATTCTAGAGCTCT 4433
QY 4146 GACACAAAGATCTCAAAACAGCAAGAGCTGTGAAA-----GCTATGAAAACGCTT 4199
DB 4434 GAGGATCATTTCACTCTAAACAGCGCTCAAGAGAGAGAAAGTTTGGAGAAAACCTCC 4493
QY 4200 GTGATTTGAGAGACAGAGAAAGTCCAGAGTGTAGGTGACACTTTATTAACAGCTGAG 4259
DB 4494 AACATTTTGAAGAACAGGTGAACGTTGAGAGCTTGAAGTGCATTTAGTTCTGAGAAAG 4553
QY 4260 AAGTCTCTGCAACGAGGTGCGCACTGACTTTTCAAGATGACAGAGACAGCGTACCCCTG 4319
DB 4554 AATCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAAAGATGCTGTGCCACA 4613
QY 4320 GGGCTGAGTCTCAGGAGCAATCCATCCCATATATATATCTCTGCTCTGAAAGCAC 4379
DB 4614 GGGCCGCACTGTGAGCAAAATGACACAGATGATATCTGCTATACCAAGAAAGGC 4673
QY 4380 CTACATCTCGACCTCAAGAGAGAAATTAAGCGATCCAGAGAGAGCGATCAGAGAGAG 4439
DB 4674 TTTAGTTCCGACTTGAAGAGAGAGAAACACATCATCTGAAGTGAAGTCAATGAATG 4733
QY 4440 GACAAGCCAGATGCTGTCTGATGCTGAACGCAAGAGAGATGACCAATCGAAAAAGTC 4499
DB 4734 GATGAGCAGGTGCTTGCC-----AGGAGGTCAAGTAGTAGTGTGCAATTTG---AGAT 4784
QY 4500 CTCAAGGCTGAACCTGAATCTGGAATTGAGAGTAAGAGCAACAATTTGTGCTGAC 4559
DB 4785 TTAGAGCTCTGAATGGAATTTTGAACCTTGAGACCAAAAGCAGTAACTTTCTCCAAAAC 4844
QY 4560 GTCATTCAGACAGCGGTGACAGTTGCGACGTAC---AGAAACAGCGCCGCAAACTCAT 4616
DB 4845 ATCATCCAGACAGCGGTGACAGTTTGTACGTACAGAAAGAACGCCAACCAAAATGTTG 4904
QY 4617 GCTTATGATTCACAGACCCAGGTTT-----CTGATGACGCTTGAACAG 4660
DB 4905 ACGTGTAGTTACACACACAGCTCACGTGATAAAGCTGACAGCGACGCTGAGACG 4964
QY 4661 CAGGAGCCCAACAGATGCTGACAAAATGAAGATGCAAGATGAAAACACCCAGTGC 4720
DB 4965 GAAACGAGAAAGAGAGAGAACCTCAGGCTCTGCAACAGATGAAAACACCAATTAAT 5024
QY 4721 GAGGCCA-----GAGAGACTTGAAGTCTGACGCTCTGGA 4759
DB 5025 TCAGCAAAAGAGAGAGTCAAGTCAACCGCATGGAACAAACATTTCTATTTTCACAA 5084
QY 4760 GGCATAGGCTCAGCTCGGAAATGCTTGCGGCTTGAAGTGAAGCGCCGAGTTCAT 4818
DB 5085 GACATGAGTGAAGCTCAGAAAAAGACATGACTGTTGAGAGTGAAGGTTCCACTGTAAT 5144
QY 4819 ----- 4818
DB 5145 GATCAGACAGCTGGAAGAGTGTCTCTCCCATCTGAGAAAGAGAGAGTGAAGCTGGAACA 5204
QY 4819 -----AGTAAGCATTTGAAGAGTGCCT 4840
DB 5205 AAGTGTGTGCCAGAAAGATGATGCTATGCTTTTGAAGCAAAAGATGAGAAATGATCATTA 5264
QY 4841 CCTCAACCCCAAGA---TCCAAAAGAGAGATGCTGTGATGAGCTCTCAGCTCCAAAGCTTA 4898
DB 5265 GTTGAACCGAAAGAGATGAAAAAGTGTATGATGTTGATGACCTCGAAGAAACAGAACTCA 5324
QY 4899 GCCCAGGAGAGGCAAGTGTCTGTGAAAACCTTAACCAAAAGAAATCCCCAGACACAAAGCA 4958

DB 5325 GCTTGGCTGATATGATGCTCTCAGAGAGCTTTAACAAAGAGTCCCAAGATACAAATGGA 5384
QY 4959 CCAAGCTTAACCGAGAGAGGAGATCCCAAAAGTTGAGGTCAGAGAAAGAAATGCT 5018
DB 5385 CCAAAACAAAAAGAGAGAGATGCCAGAGAGTGAATTTGACGAAAGGAAAAAGTGCAC 5444
QY 5019 ACCAATGATCAAAAGAAACAGAGCCAGAGACAGAAAGACCTGCGAGAGCCAAAGGA 5078
DB 5445 AGTAAATCAGTAAAGGATACACACCCAAAGCACAGAGAGATTAAGAAACAAAGAGA 5504
QY 5079 GACCTGGAGAAATCTPAAGATGTTAGTTG-----CTCATTTGAC 5117
DB 5505 GAATCTGAAGTGAAGATCTTAAGAAATCTTAAGAAATCTATGACATTAACCTCATTTGCT 5564
QY 5118 ATCTGAAGACCAAGATGTAAGAAACAAAGTCAAGAAACAA-----GATGCTGTGAGGAC 5173
DB 5565 GTTTGAAGACCAAGATGTAAGCAAGTGAAGAAAGAAATGATGCTGCTGAGAG- 5623
QY 5174 CTTGAAGACCAAGATTTCAAGAGCCATGAGATCCAGAGAGAGAGGCGCTCAATGATTTCC 5233
DB 5624 CTGAAGACCAAGTATTTGAGA-ACCTTGAGAAATGGAAGAGAGCAATCAACTGATCTCA 5682
QY 5234 ACCAGTAGAGCAACCCGCAATTTCTGAGGCTTCATGCGGAGCTAGAGAGCCAGTAACTT 5293
DB 5683 TTTCTAGAGAGC-CCCTGACAAATCTGAGGCTTCATGAGAGCTAGAGCCATTTAACTTT 5741
QY 5294 TCCTGTGTTCAAGACTGCTTGAATTTGCCCTTGATGCGGTCCGTATTTCTAACTT 5353
DB 5742 TCCTCTTTCCAAAGCAACACTCAAAATTTTCCCTTGATAC----- 5781
QY 5354 TCCTGTGTTCAAGACTGCTTGAATTTGCCCTTGATGCGGTCCGTATTTGGAATTTA 5413
DB 5782 -----CATTAATATTCGATTTA 5799
QY 5414 AGGTCTGCGCTTCAACTGGAAC-CAATTCGCACTACTAGTTCCACTTCAACT 5472
DB 5800 AGGTCTTAATTTCTTAACCTGGAACGAGTGGCAATCTTAATTTCTGCTCTGAACCT 5859
QY 5473 GAGACATCTCTCTTAATGATTTATATGATGTTTATGTAAGTCTCTCTCTGATCTAT 5532
DB 5860 GAGATATCTTTCTTAATATTTATGTAATGTTTAAAGT---TCTCTGTATCAT 5916
QY 5533 TGTATATTTTCTTAACGTTT-----AAGCAATGCTTTTGTATTA 5575
DB 5917 TGTATATTTTCTTAATGTTTAAAGAAATGTCAGAGATCTCATGCTTTTGTATCA 5976
QY 5576 TGCATATATTAACGAGGTGTGACCATAGCGAGCTTTGAAAAGCTCCAGAGCTCAACTG 5635
DB 5977 CACAGTATATGAGGAGATGTCATAGTGCAGGCTTGGAGCTTTAAAGCTCAAGTTA 6036
QY 5636 TAACCTGAGCAACAGATTAACATTC-----CTGGCAAGAGAGACAAAGTCTTTT 5686
DB 6037 TATAACCCACAAAAACAGAGCTCTAGATGTAACATTCCTGATCAAGATCAATTCCT 6096
QY 5687 TAAAGTTACTGATGCTTGAATCTGTGGGCTTGAATCTCTGAAAGTGTGTTTCTCT 5746
DB 6097 TAAATTCATCAATGATGATGAGTCCATATTTAGTGAATCTGAAATGTGTGACCTTCTCT 6156
QY 5747 ATGCACAGGAGCTCAGAAATTAACCCCATTTTGAACATCCAGAGATGCCAATTT 5806
DB 6157 ATTAACAGGAGTGTGCCAAACCTAAAAAGCTTTTGAACATACAGAAATGTTATATGTC 6216
QY 5807 AACATGATTTTTCCTCCCTTTTGTCTAATCCAGTCCAGTTGGAAGAAAGTCTCTCTG 5866
DB 6217 ATGGGAAATTTTG-----TTTCTAACAGTGAAGGTTAAGAAATTAATTTCT 6269
QY 5867 TGTAGA-----TTAAGCCCTGTCTTATATATATGACAAATGAGT 5909
DB 6270 GGTAGCAATTAATTTACATCTTTTCTTAATGTTATGTTGTTGAGCCAGTAATGT 6329
QY 5910 GTGCTTAAGGCAATGAG-----TGTTTCTAATGCAAGAAATCTGT 5953

Oy	1419	ACTAACTGACACACCTTAAGTCTCTGAAAGAAAGACGCTGGCCAAACCCCGAAAGGCATT	1478
Db	1617	ACACAGGAGACTGACCTCACTCTCTGATGAAAGTGTCGTCCAAACCCCGAAGGCCTT	1676
Oy	1479	GTCAGTAGGTGGAGATGCTGTCTCTCAAGAAAGAAATCAAGTACAGGAAAGTCCCTTG	1538
Db	1677	GTGAGTAGAGGTGGAAATAGTCTGTCATCAAGAGAGAAATGAAGTGCAGGAAGTCCACTA	1736
Oy	1539	AAGAACTCTTCAGTAGCTCAGGCTTTAAAGAGCTGTCTGGGAAGACGAAAGGGGAAA	1598
Db	1737	AAGAAGCTTTTACACACACTGGCTTTAAAAAGCTTTCTGGAAAAGAAACAGAAAGGGAAA	1796
Oy	1599	CGAGAGAGTGGGGGAGACGAAAGACCTGAGGAATACCAACATTCACACGGAATCCCCA	1658
Db	1797	AGAGGA---GGAGGAGACGAGAAATCAGGGGAGCACTCAAGTTTCAGGCGAATTCCTCG	1853
Oy	1659	GAGAGTCTGATGAGCAGAAAGGAGAGAGCTTCGTCTGCTCCCTCGAGAGACTGTAGAG	1718
Db	1854	GACAGCCAGAGGAGGCAAAAAGGGCGAGAGCTCTGCTCTCCCTGAGAGAGCCGAGAG	1913
Oy	1719	ACCACTGTCTGAGGAAAGGGCCGCTGGAAAGCACCCAGAAATGGGGAAAGCTGAGGAAGA	1778
Db	1914	ATCACGTCTCTGAAAAAGGGCTTAGCCGAGGTGCACGAGATGGGGAGCTGAAAGAA	1973
Oy	1779	ACTACTTCCGATGAGAGAAAGAAAGAGAGGAGTCACTCCCTGGGCACTCTTCAAAAAG	1838
Db	1974	GCTACTTCCGATGAGGAAAAAAGAAAGAGGTGTCACCTCCCTGGGCACTTCAAAAAG	2033
Oy	1839	ATGGTGACACCCAGAAACGGGTCCGAAACCTTCTGAAGATGACAGAGAGGAAGACTG	1898
Db	2034	ATGTGATCGCCCAAGAAACGTGTTAGACGGCTCTCGAAAGATGATTAAGAAATAGACTG	2093
Oy	1899	GAGAAAGTCAAGACGCGCACCTGTCTCTCACTGATAGACAGTGTCAAGAAATGCAAGAT	1958
Db	2094	GACAAAGTCAAGACGCTACTCTTCTTCCAGAGAGCACAGCTCTGAATGCAAGAA	2153
Oy	1959	GAAGTCAAACTGTTGTGAGAAACAAAGCCAGAGAAACCAAGCTTAGGTTGATCT	2018
Db	2154	GAATGAAAGGAGACGTGGAAGACCAAAACCCGAAAGAACCAAGGCGAAGTGAATACC	2213
Oy	2019	TCAGTCTTGGGAAAGACGATTTGTGTGGATATCCCAAGAAAGACCAAGAAAGCA	2078
Db	2214	TCAGTATTTGGGAAAGCTTTATTTTGTGTGGATCAITCAAGAAAGACCAAGAAAGG	2273
Oy	2079	TCCTCTTCAGATGATGAAAGAGGGCCCAAGGACACTGGAGGGGACAGTCAAGACAG	2138
Db	2274	TCCTCTTCTGATGAGAAAGGGGAGCAAAAAGCAATGGAGAGAGACACACAGAAACTGAT	2333
Oy	2139	GAGGCCACGAAAGACAAAGAAAGCCGGAAACGACGCTGTTCTGCCAGACACCCAGAGAG	2198
Db	2334	GAGGCCGGAAGAAAGCAAGAGAACGGGAGACAGAGGAGATCTCTGTGCTTCCCAAGAACT	2393
Oy	2139	GACCAAGGCAAGAGAGTCTCTCACCCGAGACCCAGCGGAAAGCCCTTCCGAAGGGAGAGT	2258
Db	2394	GATCCAGGGCAGGAAGTCTCTCCCGGAGACAGCTGGAAGCCCTACGAAAGGGAGGGC	2455
Oy	2259	GTCTCACTTGGGAGTCAATTTAAAGATATGCACTTCCAAAGAAAAATCCAGTCAAAA	2318
Db	2454	GTTCACCTCGGAGATCAATTTAAAGGTTAGTCAAGCCCAAGAAAAAAATCAAAAGTCCAAAG	2513
Oy	2319	CTGGAAGAGAAAGCCGAAGAC-----TCTAGGTGAGAGAGTTGTCACTGAG	2366
Db	2514	CTGGAAGAGAAAGGAGAAAGTCCATATAGCTGGGTCTGGTGTGACACATTTCACTCCAGAC	2573
Oy	2367	ATCGAAGCAGATGAGAGAAATCTTTGGGTTTTCATTTAAGAAATTCATCCCGGACGGCGG	2426
Db	2574	ACTGAACCCGGTAAAGAAAGATCTCGGGTCTCAATCAAGAAATTTATTTCTGTGACGAAGG	2633
Oy	2427	AAAGAAAGGCGAGACGGGAGCAAGAACCACTGTGGAAAGACTCAGAGGCCATGTGAG	2488
Db	2634	AAAGAAAGGCGAGATGGGAAACAAAGAACAGCCCTGTGTGAAACCCAGGGCCACAGGG	2693

Oy	2457	ATTTAATGAGACGACCTTATATGCCCCGCCGCGGCGCTCTGTCGATGATTAATGACGTG	2546
Db	2694	GCCAAACGAATATACCTGATGTGTCGGCGCGTGTCCTCTGTCTGATGATGATGCTGTGA	2753
Oy	2547	GAGAGGAGAAATGCG-----AAGCCGAGGGGAAATACGAGCTGCCCAAGCTCTG	2597
Db	2754	GAAAGGAGAAAATGAGAGCGACAGCAAGCCCAAAAAGGCGCAGAGCGAGCCGAGCACAAG	2813
Oy	2598	GGGGCTGTGTACGTGTCCGAGAGCTCAGTAAAGCTCTGTGTCACACTGTGATGTGCGCA	2657
Db	2814	GCACCCACTGAGGTGTCCAAAGAGCTCAGCGAGACTCAGGTTCAATGATGGCAGAGCT	2873
Oy	2658	GTCATTGATGGGACGAGGGGAGTCACAGTGTGCAAGAGCGGTCCCTCGTGATATACC	2717
Db	2874	GTGCTTACCGGACGAGGGGAGTACCATTTATTGAAGAAAGGTCTCCTTCTGTGATATCT	2933
Oy	2718	GCTTCCCTTAAACAACCTCTTGAACAACAGCGGGAGAGAGCATGCCACTGTGGAAG	2777
Db	2934	GCTTCACTGACAAACCTCTTGAACAAGTAAAGCTGAAGCCGACCTTTAACTGAGAG	2993
Oy	2778	GTCACTGAAAAAGACATCATTTGC--AGAGAACTCCTGTGTCCACCAGAGCTTACCA	2834
Db	2994	GTAATTGAAAGAAGATTAATGTGAGAAAGAAACCCCAACGGTTACTGAACCTCTGCCA	3053
Oy	2835	GAGGTTAAAGATGCCATGACACATGTGTACCACTGAAGTGAATTTACTCCGAAGCT	2894
Db	3054	GAGAAACGAAGAGCGCGGGCGCACAGGTCGTTAATGTAGGGGGAATTAACCCCGAAGCT	3113
Oy	2895	GTTGACAGCACAGAGACTCAGAGGCTCTCCGTACGTAAGAAAGTTACGGAAGACTCGGG	2954
Db	3114	GTGACAGTGCAGAAACTGAGGGGCCATTGGGTTCCGAAGAAAGAACCGAAGCATCTGCT	3173
Oy	2955	GCCGAAAGACCAACAGACATGTGTGTCCGACATTTCCACGTGACTGACTCTCCACAGACC	3014
Db	3174	GCTGAAGAGACACAGAAATGAGTGTGACAGTCTCCAGTTTAAACCGACTCCCAAGACACC	3233
Oy	3015	ACAGAGAAAGCACCCCACTTCAGAGAGGTGAAGAGTGTGTCTAGATACAGAAAGAGAG	3074
Db	3234	ACAGAGAGGGCACTCCGGTGCAGAGAGGTGAAGGGGGTGACTGACATGAAAGACGA	3293
Oy	3075	GAGCGCCAGAGCGCAGGCACTCCCAAGCGCGTTTGCAGACCAAGGTAAAGAGAGGTCAG	3134
Db	3294	GAGAGCGGCACTCAAGAGGTCTCTCCAGGACGTGGCAGAAAAGTAAAGAGAAATCCAG	3353
Oy	3135	GTGCTGTCAACC---CAGACTGTGCAGAGAACGGGTCAAAAGCACTGAGAAAGGTTGAG	3191
Db	3354	CTGCGCTGGCACCGGTGGGCGAAGATGTGTCTTCAAGCTGTGCAGAGACGAGGCGAGA	3413
Oy	3192	GAGGTAGAGAGAGCTCCGAAGTGTCTGCTTGGAGAAAAGAGAGAGCGTTATGTCCGAAA	3251
Db	3414	AGACACGAAAGACAGGCTGAAGCGTGGGGTCTGAAAGAAAGAACCGAGATGTAGTGTGAA	3473
Oy	3252	GGAACCCGTGCAAGAAAGTGAAGGTGAGCAATTTGCAACAGGGCTGTGAAGTGTGACAGGCT	3311
Db	3474	GTAGATGTCTGAGGGGAAAAACTGAGCCTTTTACAAAGGAGAGGTGTGTGGGCGAGCC	3533
Oy	3312	ACTCCAGAGAGCCTTGT---AAGTTCCTGAAGTACCGGACGATGTATGAC-----CAT	3359
Db	3534	AACCCAGAAAGCTTTGAAAAAGCTCTCTCAAGTACAGAGAGATAGATGTCCAGTGAAGCTT	3593
Oy	3360	GTGCGCACTGCCA-----GTTATCAAGCTCCAGACGACTGATGGA	3401
Db	3594	GTAACTTGTCAAGCCGAACCTTATAGCTGGGGTAAAAATCAACAGAAATGTGTATGANA	3653
Oy	3402	CAGGCGGTGGCCCTGTAGTCACTCGAAACCTTGAACGAAAGTGAACAATAATGGAAGCAT	3461
Db	3654	CAGGCTATCCCCCTGTGCTGTGTGAAACCCCTTACGACGATGGAACCTGTATGGAAGCAC	3713
Oy	3462	CCCTTAGAGATTAGACACTGACAGATGGGACACAGCAAGATGAAGAACATTGACAGCCAG	3521
Db	3714	CCCTGAGCCGACTTTGACGACACAGGACCAACCCAGAAAAGAGAGATTTGTGAAATTCAT	3773
Oy	3522	GACATTAAGCACTGCAGCTGTTCAGGACGTCAAGTCAAGAAAGAGAGCGGCTACT	3581

Db 3774 GAGGAGATGAGTGCATCTGGTACCGAGTCAGGGGCAACGAGAGAGAGTTCCT 3833
Qy 3582 GCTCAGAAAGAGAGCCTTCGACACTACCTAATATGTTCCAGCCCAAGGAACATGGG 3641
Db 3834 GCACAGAAAGAGAGCCTCCAGCACCTTCCAGTTTGTGTTCCAGGAAGAACTAAAGAA 3893
Qy 3642 GAAGAACCGAGAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGTGACGCGTG 3701
Db 3894 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTCAAGTGAACCTGTA 3953
Qy 3702 CCCCTTCGCAAGACTGAGGTGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3954 TCCATTCTGTCAAGACTGAGGGACTCAAGAGGCTGACCGATATGCTGATGAGAAAC 4013
Qy 3753 GATGAGAAAAAGTCAAGAGAAACAGGAGGTGTTGTATACACTCTGG-----ACCCAAC 3806
Db 4014 AAGACGTACCATTTTTCGAAGGACTTGAAGGGTCTATAGACACAGGCATAAACAGTCAGT 4073
Qy 3807 AGTCAAAAGGCTGTGATGTGACATATGACAGTGAAGTATGAGGAGTGGCGGTGTGAG 3866
Db 4074 CGGAAAAGGTCACTGAAGTTGGCCCTTAAAGGTGAAGGGACAGAAAGCTGAATGTAAA 4133
Qy 3867 GAAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGGAGGA 3911
Db 4134 AAGGATGATGCTCTTGAACCTGACAGGTCAAGTAAAGTCTCTCCATCCCCCGTGGAGAGA 4193
Qy 3912 GAGATGGAAGCTGAGTTGAAAAGGAGAAAGGAGAGACAAAGCCAGAGCAAGTGAAGTAA 3971
Db 4194 GAGATGTTAGTTCAAGTCGAAGGAGAGAAACAGAGCAGAGCCAAACCCATGTGAATGAA 4253
Qy 3972 GAAGG---TGAGCAGGAAACAGCGCTCTGAGCATGAAGGAACCTACGGGAAGCCAGTC 4028
Db 4254 GAGAAGCTTGAGCAGAAACAGCTGTTACCGTATCTGAAGAGGTCAAGTAAGCAGCTCCTC 4313
Qy 4029 CTGACACTGACATGCCAGCTCAGAGGGGGAGGACACTGGGAAGCCTTGGAGGAAGC 4088
Db 4314 CAGACAGTGAATGTGCCCATCATAGATGGGGCAAGGAAGTCAGCAGTTTGGAAAGGAAGC 4373
Qy 4089 CTTTCTC---TCCAGACCAAGACAAAGAGCGTTGCAATAGAGGTTCAAGTTCAAGGCTG 4145
Db 4374 CTTCTCCTCGCTAGGTCAAGAGGAGCAGTATGCACCAAAATTCAGTTTCAGAGCTCT 4433
Qy 4146 GACACACAGTCACTCAACAGCAGAGAGCTGTGAAAA-----GGTCATAGAAACCGTT 4199
Db 4434 GAGGCATCAITTCATCTTAACAGCGGCTGCAGAGGAGAAAGGTCTTAGGAGAACTGCC 4493
Qy 4200 GTGATTTTCAGAGCAGGTGAAGTCCAGAGTGTGTAGGTGCACACTTATACAGCTGAG 4259
Db 4494 AACATTTTAGAAAACAGGTGAACCGTTGGAGCCTGCAGGTGCACATTTAGTTCTGGAAGAG 4553
Qy 4260 AAGTCTCTGCAACCGGTGGCCACTGGACTCTTCAGCATGCAGAGGACACGGTACCCCTG 4319
Db 4554 AATCCTCTGAATAAATGAAGACTTTGGCCCTCATCCAGGGGAAGATGCTGTGCCACA 4613
Qy 4320 GGGCCTGAGTCTCAGGAGAAATCCATCCCAATAGTAATCTCTGCTCCTCGAAAGACC 4379
Db 4614 GGGCCGAGTCTCAGGCAAAATCGACACCAGTATAGTATCTGCTACTACCAAGAAAGGC 4673
Qy 4380 CTACATCTGACCTACAGGAGAAATAGCCATCCAGAGAGAGCCATCAGAGGAAGAG 4439
Db 4674 TTAAGTTCCGACCTGGAAGGAGAGAAACCACTCATCTGAAGTGAAGTCAAGTGAAGTC 4733
Qy 4440 GACAAGCCAGATGCTGTCTGTGCTGAGCGGCAAGGAGAGTACGCAATCGAAAAAGTC 4499
Db 4734 GATGAGCAGGTGCTGTGCC-----AGGAGTCAAGTGAAGTGTAGCAATG---AGAT 4784
Qy 4500 CTCAGGCTGAACCTGAGATCCTGGAATCTGAGAGTGAAGAGCAACAAGATTGTCTGAAC 4559
Db 4785 TTAGAGCTGAATAATGGATTTCGAACTTGAGACCAAGAGCAGTAACTTGTCCAAAC 4844
Qy 4560 GTCAATTCAGACAGCGGTGACAGTTCGCAGTAC---AGAAACAGCCCCCGAAACTCAT 4616
|||||

Db 4845 ATCATCCAGACAGCCGTTGACCAGTTTGTACGTACAGAGAAACAGCCACCCGAAATGTTG 4904
Qy 4617 GCTTATGATTTACAGACCCAGGTTCT-----CTGCAATGCAAGGTTGACAG 4660
Db 4905 ACCTCTGAGTTACAGACACAAGCTCACGTGATAAAAGCTGACAGCCAGGACGCTGGACAG 4964
Qy 4661 CAGGGAGCCCAACAGATGCTGGACAAAATGAAAGATGCCAAGATGAAACACCCAGTGCC 4720
Db 4965 GAAACGGAGAAAGAGAGAGGAAACCTCAGGCCCTCTGCACAGATGAAACACCAATTAAT 5024
Qy 4721 GCAGGCCA-----GAGAGGACTTTCGAAGTCTGCAAGTCTGACCGTTCTGGA 4759
Db 5025 TCAGCCAAAGAGAGAGTCAGAGTCAACGGCAGTGGGAACAGCACATCTGTATATTTCCAAA 5084
Qy 4760 GGCAATGGGCTCAGCCTCGGAAATGCTTGGCCGCTTTCAGATTTGAAAGCGCGGTGTCAA- 4818
Db 5085 GACATGAGTGAAGCCTCAGAAAAGACCATGACTGTTTGAAGTAGAAGTTCACATGTAAT 5144
Qy 4819 -----4818
Db 5145 GATCAGCAGCTGAAGAGGTGCTCTCCCATCTGAGNAAGAGGAGGTGGAGCTGGAACA 5204
Qy 4819 -----AGTAAGCATTTGAGAAGCTGSCCT 4840
Db 5205 AAGTCTGTGCCAAGATGATGTCATGCTTGTAGCAGAAAAGATAGAGAAGTCACTA 5264
Qy 4841 CCTCAACCCCAAGA--TCCAAAAGAGAGCATGCTGATGGCCCTCAGCTTCCAAAGCTTA 4898
Db 5265 GTTGAACCGAAGAGATGAANAAGGTGATGATGTTGATGACCTTGAANAACAGACTCA 5324
Qy 4899 GCCAGCAGAGGCGCAGTCTCTGGAACACTTAAACAAAGAAATTCCTCCAGACACAACGGA 4958
Db 5325 GCCTCGCTGATCTGATGCTCAGGAGGCTTAAACCAAGAGTCCCCAGATACAAATGGA 5384
Qy 4959 CCAAGCTTAACCCAGAGGCGGATCCCAAAAGTTGAGGTCAGGAAGAGAAATGTCT 5018
Db 5385 CCAAAACAAAAGAGAGAGGAGTCCAGGAATGAGAAATTGCAGGAAGGAAAAGTGCAC 5444
Qy 5019 ACCAAGTCAGTCAAGAGAAACAAAGGCCAGGCAAGAGAGACCTGCAGGAGCCAAAGGGA 5078
Db 5445 AGTGAATCAGATAAGCGATCACACCCCAAGCAGAGGAGGTTACAGAAACAGAGAGA 5504
Qy 5079 GACCTGGCAGAAATCTTAAGATGTTAGTTG-----CTCATGTATC 5117
Db 5505 GAATCTGCAAGTCAGAACTTACAGAACTCTTAAACATCATGCAGTTAAACTCATGTCT 5564
Qy 5118 ATCTGTAAGACCAAGATGTGAACCAAGTCACAGACAA-----GATCCTGCTGTGGAC 5173
Db 5565 GTTGAAGAGCCAGAAATGTGAAGACAAGTAGTAGAAGAAATGAATGCTGCTGTGAGA- 5623
Qy 5174 CTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGCAGGGCGTCCAAATGATTTC 5233
Db 5624 CTGAAGCCAGTATTTTCAGA-ACCTTGAAGATTTGGAGAGCAGGCACATCAACTGATCTCA 5682
Qy 5234 ACCAGTAGAGACACCCGACAAATTCAGAGCTTCATCGGAGCTAGAGCCAGCTAACATT 5293
Db 5683 TTTCTAGAGAGC-CCCTGACAAATCCTGAGGCTTCATCAGGAGCTAGAGCCATTAACTT 5741
Qy 5294 TCCTCGTTTCAGACTGCTTTGATTTGCCCTTGATGCGCTCGGTATTTCTTAACATT 5353
Db 5742 TCCTCTTTCCAGACCAACCTACAAATTTTCCCTTGTAAAC-----5781
Qy 5354 TCCTCGTTTCAGACTGCTTTGATTTGCCCTTGTGATGTCGCTGCTGATTTTCGGATTTA 5413
Db 5782 -----CATATAAATTCGATTTA 5799
Qy 5414 AGTCTCTGCTTCTCAACCTGGAAC-CAATTCGCCATACCTAGTTCCACTTCTCAAACT 5472
Db 5800 AGTCTCTAAATCTTAACCTGGAACCTGGAGTTGCAATACCTAGTTCTGCTTCTGAACT 5859
Qy 5473 GGAGCATCTCTTTATGATTTATATGATGATGTTTTATGATGCTCTCTCTCTGCTGAT 5532
Db 5860 GGAGTATCATCTTTTACATATTTATATGATGTTTTAAAGTAG---TCCTCCTGTATCTAT 5916
|||||

QY 5533 TGTATATTTTCTTACGCTT-----AGCACATGCTTTGTATTA 5575
 Db 5917 TGTATATTTTCTTATGTTTAAAGAAATGTGACAGATCTACTACTCTTTTGTATCA 5976
 QY 5576 TGCATATATTAACGGGTGTGACATGACGAGCTTTGAAAGCTCCAGAGCTCACTG 5635
 Db 5977 CACAGTATATATGAGGATGTCATGTCATGTCAGGCTTTGAGGCTTAAAGCTTACGTTA 6036
 QY 5636 TAACCTGACGAAACAGATTAACATTC-----CTGGCAAGAAAGACAAAGTCTTTT 5686
 Db 6037 TATATACCAAAAAGAGAGCTCTAGATGTATACATTCCTGATCAAGTACATTTCT 6096
 QY 5687 TAAAGTTACTGATGCTTAAATCTGTGGGCTTCTGCTCTGAAAGGTGTTTCTTCT 5746
 Db 6097 TAAATTCATTAATGATGAGGTCCATATTTAGTGTACTCTGAAATTTGCTACCTTCT 6156
 QY 5747 ATGCACAGAGCTCAGAAATTAACCCATTTTGAACATCCAGATGTCCTCAATAT 5806
 Db 6157 ATTACACGAGAGTGTCCAAACTAAAGACATTTGAAACATACAGAAATGTTCTTATGTC 6216
 QY 5807 ACCATGATTTTCCCTTTTCTTATCCAGTCCAGGTGGAAGAGTCTCTCTG 5866
 Db 6217 ATTGGGAATTTTGC-----TTTCTAACCCAGTGGAGTTAGAAAAGATATATTTCT 6269
 QY 5867 TGTCTAGA-----TTAAGCCCTGTCTCTTAATGATATGACAAATAGT 5909
 Db 6270 GGTACCAATTAACCTTACATCTTTTCTTACTTGTATGTTTGGACCGAATAGT 6329
 QY 5910 GTGGCTTAAGCCATGAGA-----TGTTCCATAGCGAAGAAATCTGT 5953
 Db 6330 GTGCTTAATTCGAGAGCAAGTAGTGAATATGTTTATATGTTAAGAAAAGAAATTTGT 6389
 QY 5954 TGTACGTTTTTGTGATGCTCTTCTATGCTGACCGAATTCATATGACATGCAATGAGTG 6013
 Db 6390 TGTATAG--TTTTGATTTACTCTTATATGCTGAGATGCTATCAACATGATGATGAATA 6447
 QY 6014 AGTCTGTTCTTAAAGATGATTTTGTATGATATGATATGATGATGATGATGATGATGATG 6073
 Db 6448 AGTACGATCTCTTAAAGATGATTTTGTATGATGATGATGATGATGATGATGATGATGATG 6507
 QY 6074 TGCCCTCTTCTTAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 6133
 Db 6508 TGCCATCTCTTAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 6567
 QY 6134 CTGATTTAAATAAA 6147
 Db 6568 CTGATTTAAATAAA 6581
 RESULT 11
 HSU81607 6608 bp mRNA linear PRI 26-JUN-1997
 LOCUS HSU81607
 DEFINITION Homo sapiens gravin mRNA, complete cds.
 ACCESSION U81607
 VERSION U81607.1 GI:2218076
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 6608)
 Naureit, J.B., Klauck, T.M., Langeberg, L.K. and Scott, J.D.
 Gravin, an autoantigen recognized by serum from myasthenia gravis
 patients, is a kinase scaffold protein
 Curr. Biol. 7 (1), 52-62 (1997)
 JOURNAL
 MEDLINE 97153077
 PUBMED 9000000
 2 (bases 1 to 6608)
 Naureit, J.B., Klauck, T.K., Langeberg, L.K. and Scott, J.D.
 Direct Submission
 Submitted (10-DEC-1996) Biochemistry, OHSU, 3181 SW Sam Jackson
 Parkway, Portland, OR 97201, USA

COMMENT On Jun 26, 1997 this sequence version replaced gi:1850362.
 FEATURES
 source
 1..6608
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 192..5537
 /note="kinase scaffold protein"
 /codon_start=1
 /product="gravin"
 /protein_id="AAC51366.1"
 /db_xref="GI:2218077"
 /translation="MGAGSSTEQRSPEQPEGSSTPAEPEPSGGSPSAENAPDTADP
 ALASPATLQKNGSLSTINGVADDELSTLOEGDNLQKNGALNGLNGSEEV
 IYVWGORDSEVSRSDSKEMATKSAVVHIDITDQGFERNRNEIOLPSESNLELTO
 PTSGQNDIGFKVYKPTVVKDKTKETPTVTLTKYKGGGAGAGHOPPS
 LGGEAASKESEPKQSTERPEETLKQESHAETISPAESQAAVEBCKEBOEKOEP
 SKASBPSPVSTSETGTFKPKPTQGMAMKRTSRPKPEDEVEASEKKKEQEPKV
 DTEDEDAEVAESKLTASDAQHPQEAEMHBERLSEAEKVELPSEBQVSGQPSB
 EKPAPLATEVDFDEKIVHOEVVAEVAEVEERTEOEKTEVETAGSVPAEVLGMD
 AEPOEAPPAKEIVKIKETCVSGEDPTQADLSPDEKVLSPKPGVSVEMMLSSORM
 KVGSPKLTFTSTGLKSLSKKKOKKRGGDESEBHTOVPAIDSDSOEQGESBA
 SPSEPEEITCLEGLAEVQDGEAEAGTSDSEKREBGTVPMASTFKMVTPKKVR
 PSEBDEBELDKYSATLSTESTASBOMEAMSVESVPEPKKRVTPVSEWELI
 CVGSSKKRARSSDEEGPKAMGDDHQADAGDKETGDIILAGSEHDPGGS
 SPSEOGSPTEGEGVSTWESFKLVTPRKSKSLKESKSDS IAGSGVSEHTDTEPG
 KESWVSIKKFTPRRKRPDQKQKQAPVEDAPGANGEDSDVPVAVPLSEYDAVER
 EKAEAOQAGAPPOPKATVEYSKLSBSQVHMMAAVALDTRATIIIEKSPNIS
 ASVTEPLEQYEAALALTELEVERVIAEEBPTVEPPEENEAAGDVTSABELTP
 EAVTAETAPLPGSEBGTASAEETEMVSAVSOLDSPTEBATEPVOEGVGD
 IEOERTQEVLAVAEKVESOLPGTGPEVLOPVQAEAREPEBOAEAGLKE
 TDVVLKVDQAEATETPTQGVVGGTTPSEFEKAPQVTEISESELVTCQATLNGV
 KSDMEMOCHLPPDSVETPTDSTDSPTPADPADGATQKDEIVIEHENEVAASGO
 SGTBEAIVNAQKERPPAPSSPVFOERTKQSMEDTLETIDKESVETYSILSKTBG
 SGTADQVADKTXDVPPEBGLBSIDGLTVSEKTEVALKQEGTEABCKDDALE
 LQSHAKSPSPVEREMVQVEREKTEAPHPHVEKLEHETAVTSEVSKOLLQTN
 VPLIDAEKVSLESGPPCLGOEAVCTKIQVSSASEFTLAAEEVEVLSBTAKI
 LEETELPEAPGALVLEESSEKNEBPAGSDAVPTGPDCAKSPVPAATTKG
 LSSDLBERKTTSLKMSDVEDVOAVQKVSVAIDLEPENLILLETESKLQNI
 IQTAVDQFVTEETATBMLTSELQTAHVIKASODAGOTEBEGBEPQASODEFTI
 TSKAESESTAVQASHDISKDMSEASEKMTVEVSGTIVNDQLEEVLPSEBEGG
 AGTQSVDEDDHALAERIEKSLVERPEDEKGDVDDPENONSAALADTDSGGITKXS
 PDTNGPFOKKEBAOEVELOEGKVSHESDVAITPQAOEILQKERESEAKSELTES"
 BASE COUNT 2091 a 1418 c 1827 g 1272 t
 ORIGIN
 Query Match 34.8%; Score 2144.4; DB 9; Length 6608;
 Best Local Similarity 64.3%; Pred. No. 0;
 Matches 4215; Conservative 0; Mismatches 1796; Indels 543; Gaps 45;

QY 309 GAGTCAGAAGATGTGAGAGAAAGACCGAGTTTGAAGAAATGGCGCCCAACTCCACAGCT 368
Db 522 GACTCTGAAGATGTGAGCGAAAGAGACTCCGATAAAGAGATGGCTACTAAAGTCAGCGGTT 581
QY 369 GTTGAAGATATCAAAAGGATGGCGAGGAGAGACATCAGAAATAAATTGAACAGATCCCT 428
Db 582 GTTCACGACATCACAGATGATGGCGAGGAGGAGA--ACCGAATATCGAACAGATTCCT 638
QY 429 GCTTCAGAAACAATGTGAAGAAATGGTACAGCTGTGAGTCCAGGCTTAATGATGTT 488
Db 639 TCTTCAGAAAGCAATTTAGAAAGACTAACAACACCACCTGAGTCCAGGCTTAATGATATT 698
QY 489 GGCTTCAAGAAAGATTTAAATTTGTTGTTTAAATTTACGGTGAAGAGGATAAAAT 548
Db 699 GGATTTAAGAAAGGTTTAAAGTTTGGCTTTAAATTTCACTGTGAAGAAAGGATAAGACA 758
QY 549 GAAAGTCAGATACTCTCAACTACTCACTGTCAAGAGGATGAAGCGAAGGGGAGAA 608
Db 759 GAGAAGCCTGACACTGTCCAGCTACTCACTGTGAAGAAAGATGAAGGGGAGGAGCAGCA 818
QY 609 GCCTCTGTGGAGTGGAGACCAAGAGCCAGTGTGAGACTGCGCTCGGAGATCA 668
Db 819 -----GGGGCTGGCGACCAACAGGACCCAGCCCTTGGGGCT-----GGAGAAGCA 863
QY 669 GCATCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGCAAGAGGACCCCTGAAGCAA 728
Db 864 GCATCCAAAGAAAGCAAGCAACCAATCTACAGAGAAACCCGAAGAGACCCCTGAAGCGT 923
QY 729 GAAAGAGCAGCAGCAAGAAATCCCCCTTCAAGCCGAATCTGATCAAGCGGCTGAGGAAGAA 788
Db 924 GAGCAAGCCAGCAGAAATTTCTCCCCCAGCCGAATCTGGCCNAGCAGT----GGAGGAA 980
QY 789 GCCAAGATGAAGAGAGAAAGAAACAAGAGAAAGAGCCCAAGTCCCAAGATCCCG 848
Db 981 TGCAAGAGAGGAGAGAGAGAAACAAGAAAGAAAGAAAGAAAGTCTGAGAAATCTCG 1040
QY 849 AGCAGCCAGTCAACAGTGAAGCAACATCTCTTCAAGAGTCTTCACTCAGGTTGG 908
Db 1041 ACTAGTCCCGTGACAGTGAAGCAAGGATCAACCTTCAAAATAATTTCTCACTCAGGTTGG 1100
QY 909 GCGGCTGGCGCAAGAGACAGCTTCAAGAAATCAAAAGAGGATGATCTGGAATCTGCC 968
Db 1101 GCGGCTGGCGCAAGAGACAGTTCAGAGAGCGGAGGAGATGAAGTGAAGCTTCA 1160
QY 969 GAGAAGAAAGGAGCAAGAGGAGCAAGAAAGTGAAGAGAGAAAGAAAGAAAGAGAG 1028
Db 1161 GAGAAGAAAGGAGCAAGAGGAGCAAGAAAGTGAAGAGAGAAAGAGAAAGAGAGAG 1220
QY 1029 CCAGCTCGGAGG-----GAGAGAGCGGCAAGAGAC 1061
Db 1221 GTTGCTCCGAGAAACTGACCGCTCCGAGCAAGCCCAACACAGGAGCGCGCAAGAAAGT 1280
QY 1062 ACAGACCAGGCGAGTGTGTCAGAGACTACAGAGAGTGGAGCTGCCTTTGGAGAGCAG 1121
Db 1281 GCCACAGGCGCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGAGCAA 1340
QY 1122 GTTGGTCACTGGAGGATCGTCAGAGAGAGAGTGTCTCTTTGGCAACGGAAGTGT 1181
Db 1341 GTCAGTGGCTCGCAGGACCTTCTGAAGAGAAACCTGCTCGTGGCGACAGAAAGTGT 1400
QY 1182 GATGAGAAAGTGGAGCCCAAGAGAA-----GTTGTCAGAGGTCACGTCAGACCCGTG 1238
Db 1401 GATGAGAAATAGAAAGTCCACCAAGAGAGGTTGTGGCCGGAAGTCCACGTCAGCACCGTG 1460
QY 1239 GAGAAGACAGAGGAGGAGGAGGAGGAGGAGGCTGAAGGGGCGTGGTGTAGAA 1298
Db 1461 GAGGAGAAACCGAGAGCAGAA-----ACGGAGGTGGAA 1496
QY 1299 GGAAACAGGAGAAATCTTTCCCGCTTGAAGAACTGGCTGAGCCCGCAGGAGTCCCGCAGGAA 1358
Db 1497 GAAACACAGCGGTCTGTGCCAGCTGAAGAAATTTGGTTGAATGGATGACAGAACTCAGGAA 1556
QY 1359 GCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGAGGAGAGCCAC 1418

Db 1557 GCCGAACCTGCCAAGGAGCTGGTGAAGCTCAAAAGAAACGTGTGTTTCCGGAGAGGCCCT 1616
QY 1419 ACTCAACTGACAGACACCTTAAGTCTCTGAAGAGAACGCTGCCAAAACAACCCAGAGGCAATT 1478
Db 1617 ACACAGGAGCTGACCTCAGTCTGATGAGAAGTGTGTCTCAAAACCCCGGAGGCGTT 1676
QY 1479 GTCAGTGAGTGGAGATGCTGTCTCTCAGGAAAGAAATCAAGTACAGGAAAGTCCCTTG 1538
Db 1677 GTGAGTGAGTGGAAATGCTGTCTATCACAGGAGAGAAATGAAGTGCAGGAAAGTCCACTA 1736
QY 1539 ACGAATCTTTCAGTACCTCAGGCTTAAAGAGCTCTCTGGGAGAGCAGAGGGGAAA 1598
Db 1737 AAGAAGCTTTTACCGAGCACTGGCTTTAAAGAGCTTCTTGGAAAGAAACAGAAAGGAAA 1796
QY 1599 CGAGGAGTGGGGGAGACGAAGAGCTTGAGAAATACCAACACATTTCAACCGAATCCCCA 1658
Db 1797 AGAGGA--GGAGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATTTCTCG 1853
QY 1659 GAGAGTGTGATGAGCAGAAAGGAGAGAGCTGTGCTGTCTCTCCCGAGGAGCCTGAGGAG 1718
Db 1854 GACAGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCCTCATCTCCCTGAGGAGCCGAGGAG 1913
QY 1719 ACCACGTGTCTGAGAAAGGGCGCTTGAAGACACCCAGAGTGGGAGAGCTGAGGAAGGA 1778
Db 1914 ATCACGTGTCTGAAAGGGCTTTAGCCGAGGTGCACAGAGTGGGAAAGCTGAAGAAGGA 1973
QY 1779 ACTACTTCCGATCGAGAGAAAGAGAGAGGAGTCACTCCCTGGGCATCTCTTCAAAAAG 1838
Db 1974 GCTACTTCCGATCGAGAGAAAGAAAGAGAGAGTGTCACTCCCTGGGCATCATTTCAAAAAG 2033
QY 1839 ATCGTCAACCCAAAGAAACGGGTCCGAAAGACCTTTCTGAGAGTCAAGAGGAGAGAGCTG 1898
Db 2034 ATGGTGACGCCCAAGAAAGCGTGTAGACGGCTTTCGAAAGTGAAGAGAGATGAGCTG 2093
QY 1899 GAGAAAGTCAAGAGCCCACTTGTCTCTCACTGATAGACAGTGTCAAGAAATGCAAGAT 1958
Db 2094 GACAAGTCAAGAGCGCTACCTTGTCTTCCACGAGAGCACAGCTCTGAAATGCAAGAA 2153
QY 1959 GAAGTCAAAAATCTTGGTGAAGAAACAAAGCCAGAGAGAAACAAAGCGTAGGGTGATCT 2018
Db 2154 GAATCAAGAGGAGCTGGAAGAGCCAAAGCCGAGAGAACAAAGCCGAGGTGGATACC 2213
QY 2019 TCAGTCTTTGGAGACACTGATTTTGTGTGGATCATCAAGAGAGAGAGCAAGGAAGGCA 2078
Db 2214 TCAGTATCTGGGAAGCTTTAATTTGTGTGGATCATCAAGAGAAAGAGCAAGAGAGG 2273
QY 2079 TCCTCTTCAGATGATGAAGAGGCGCAAGGACACTGGGAGGAGACAGTCAAGAGCAGAG 2138
Db 2274 TCCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGGAGGAGACCAACAGAAAGCTGAT 2333
QY 2139 GAGGCCAGCAAGACAAAGAACCGGAAACAGAGCTGTTCTCCAGCAGCAGGAGGAGCAG 2198
Db 2334 GAGGCCGGAAGAAAGACAAAGAGACGGGGAGCAGAGGAGTCTTGTCTGTTCCCAAGAACAT 2393
QY 2199 GACCAAGCCGCAAGAAAGTTCTCTCACCCGAGCAGCGGGAAGCCCTTCCGAAGGGGAGGT 2258
Db 2394 GATCCAGGCGAGGAGTTCTCTCCCGAGAGAGCTGGAAGCCCTACCGAGGGGAGGC 2453
QY 2259 GTCTCCACTTTGGAGTCAATTTAAAGATTAGTCACTTCAAGAAAGAAATTCGAAGTCAAAA 2318
Db 2454 GTTTCACCTCTGGAGTCAATTTAAAGGTTAGTCAACGCAAGAAAGAAATCAAAAGTCCAAG 2513
QY 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTGTCTCACTGAG 2366
Db 2514 CTGGAAGAGAAAGCCGAAGACTCCATAGCTGGGTCTGGGTGAGAACTTCCACTCCAGAG 2573
QY 2367 ATCGAAACCGAGTAGAGAGAAATCTTGGGTTTCCATTAAGAAATTCATCCCGGAGCGCG 2426
Db 2574 ACTGAACCCGGTAAAGAGAAATCTTGGGTCTCAATCAAGAGTTTATTTCTTGACGAGG 2633
QY 2427 AGAAAGGGGAGACGGGAGGAGCAAGCAAGCCACTGTGGAAGACTCAGGGCCAGTGGAG 2486

Dh 2634 AAGAAAAGCCAGATGGGAACAAGAACAGGCCCTGTTGAAGACGAGGGCAACAGGG 2693
Qy 2487 ATAAATGAGACGACCTTAATGTCCAGCCGTGGTGGCTTGTGTGATATAATGCACTG 2546
Dh 2694 GCCAACGAAATGACTGTGATGTCGGCCGGTGTGCTCTGTGTGATATGATGCTGTA 2753
Qy 2547 GAGAGGGAGAAATGAG-----AAGCCAGGGGAATACGGAGCTGCCAGCTGTG 2597
Dh 2754 GAAAGGGGAAATGAGAGCAGCAGAGCCCAAAAAGGCGAGAGCAGCCGAGAGAAAG 2813
Qy 2598 GGGGCTGTGATGCTGTCCGAGAGGCTCACTAAGACTGTGTCCACACTGTAGTGTGCA 2657
Dh 2814 GCAGCACTGAGTGTCCAGAGAGCTCAGCAGAGTCAAGTTATATGATGGACAGACT 2873
Qy 2658 GTCAATTGAGGACCAAGGACGCTCACAGTGTCAAGAGCCGTCTCTTCTGTGATATCC 2717
Dh 2874 GTGCTGACGGGACGAGGGGAGCTACCATATTTGAAGAAAGGTCTCTTCTGTGATATCT 2933
Qy 2718 GCTTCCGTAACAGAACCTTTGAACACACAGGGGAGAAAGCATGSCACCTGTGAGAG 2777
Dh 2934 GCTTAGTGAAGAACCTTTGAACAGTGAAGCTGAAGCCGCACTGTTAACTGAGAG 2993
Qy 2778 GTCACTGAAAAAGACATATTGCT---AGAAAGAACTCTGTGTCTCACCCAGAGCTTTACCA 2834
Dh 2994 GTATTGAAAAGAGATATTGTCAGAAAGAAACCCCAAGGTTACTGAACTCTGCA 3053
Qy 2835 GAGGTTAAAGATCCCATGACGACATGTGTCACCACTGAAGTGAATTTCACTCAGAGCT 2894
Dh 3054 GAGAACAGAGAGGCCCGGGGCGACACGCTGCTTAGTGAAGCGAAATTTGACCCCAAGCT 3113
Qy 2895 GTGAAGCCACAGAGACCTCAGAGGCTCCGTACTGAAGAAAGTTACCGAAGCATCGGG 2954
Dh 3114 GTGACAGCTGCGAAGAACTGCAAGGCGCATTTGGTTCGAAAGAAAGAACGAACTGTGT 3173
Qy 2955 GCCGAAGAGACCAAGACATGTGTCCGCAAGTTTCCAGCTGACTGATCTCCCAAGACC 3014
Dh 3174 GCTGAAGAGACCAAGAAATGTGTAGCAGTCTCCAGTTAAACGATCCCAAGACACC 3233
Qy 3015 ACAGAGAAAGCAACCCCACTTCAAGAGTGAAGAGTGTGTGTCTAGATACAGAAAGAG 3074
Dh 3234 ACAGAGAGGCGCACTCCGCTGAGAGGTGGAAGGTGGGTACTGACATPAAGAAAGCAA 3293
Qy 3075 GAGCCCCAGACCCAGGCCATCTCTCAAGCCGTTGACAGCAAGAGTGAAGAGAGTCCAG 3134
Dh 3294 GAGAGCGGACCTCAAGAGGTCTCTCAAGGAGTGGCAGAAAGTGAAGAAAGATCCAG 3353
Qy 3135 GTGCTGCAACCC---CAGACTGTGACAGAAAGGGGTCAAAAGCACTGAGAAAGTTGAG 3191
Dh 3354 CTGCTTGGCACTGGTGGGCGCAAGATGTGCTTCAAGCTGTGACAGACAGAGGAGAA 3413
Qy 3192 GAGGTAGAGAGAGACTCCGAAGTGTGCTTGGAGAAAGAGAGAGCTTATGCCGAAA 3251
Dh 3414 AACACGAAAGACAGGCTGAAGGTGGGTCTGAAGAAAGAGACGATATGATGTTGAAA 3473
Qy 3252 GGAACCGTGCAGAAAGTGAAGTGAAGCTTTGACAGAGGCTCTGAGCTGAGAGGCT 3311
Dh 3474 GTAGATGCTCAGAGGCAAAAAGTGAAGCTTTTACAGAAAGAGGTGTGGGGCAGACC 3533
Qy 3312 ACTCCAGAGAGCTTGT---AAGTTCTGAAGTCAACGGCAGATGTAGAC-----CAT 3359
Dh 3534 ACCCCGAAAGCTTTGAAAAAGTCTCTCAAGTCAAGAGAGCTTAAAGTCCAGTAGCTT 3593
Qy 3360 GTGCGCAAGTGTCCA-----GTTATCAAGTCCAGCAGCTGATGAAA 3401
Dh 3594 GTAAACCACTTGTCAAGCCGAAACCTTAGCTGGGGTAAAAATCACAGAGATGTGATGAAA 3653
Qy 3402 CAGGCGGTGGCCCTGAGTATCCGAACCTTGAAGAGAGTGAAGCAATGGAAGCACT 3461
Dh 3654 CAGGCTATATCCCTCTGACTGTGTGAAACCTTACAGACAGTGAAGTGAAGGACACC 3713
Qy 3462 CCCTTAGAGATTCAGACACTGAGATGGGACACAGCAAGATGAACCAATTGACAGCCAG 3521
Dh 3714 CCGTAGCCGACTTTGACGACCAAGCAACCCAGAAAGAGAGATTTGTGAAATCCAT 3773

Qy 3522 GACAGTAAAGCCCACTGAGCTGTCAAGGCACTCAAGGTCAAGAAAGAGAGGCGCTACT 3581
Dh 3774 GAGAGAAATAGAGTGTGATCTGGATCCCACTCAGGGGGCACAGAGAGGCAAGTTCT 3833
Qy 3582 GCTCAGAAAGAGAGCCTTGTGACACTACCTAATATATGTTCCAGCCAGAAAGCATGG 3641
Dh 3834 GCACAGAAAGAGAGGCTCCAGACCTTCCAGTTTGTGTTCCAGAGAAAGCAATAAGAA 3893
Qy 3642 GAAGAACCCAGAAAGATGTTCTTGAACCTTACAGCAAGAGCTTACTGTGACCGCTG 3701
Dh 3894 CAATCAAAAGATGAAGCACTTGAAGCATACAGATPAAGAGGTGTAGTGAAGAACTGTA 3953
Qy 3702 CCGGTTCTGGCAAAAGACTGAGTGTGGTCAAGAGGCTGA-----GTTTACTGTGTTG 3752
Dh 3954 TCCATTCTGTCAAAAGACTGAGGGGACTCAAGAGGCTGACAGATGTGTATGAAAAACC 4013
Qy 3753 GATGAGAAAAAGTCAAAAGAAACAGAGAGTGTGTTGACACTGTGG-----ACCPAAC 3806
Dh 4014 AAAGACGTACATTTTTCGAAGGACTTGAAGGGGTCTATAGACACAGGCATTAACAGTCAGT 4073
Qy 3807 AGTCAAAAGGCTGTGATGTGACATATGACATGTAAGTGAATGGAGTGGCCGGGTGTGAG 3866
Dh 4074 CGGAAAAAGTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAAAGACTGAATGTAAA 4133
Qy 3867 GAAAAGGAGATGACTGAAGTGTGACAG-----TCTTAAGCTGTGAGAGAGGA 3911
Dh 4134 AAGATGATGTCTTGAATCTGACAGATCACGCTAAGTCTCTCCATCCCGGTGAGAGA 4193
Qy 3912 GAGATGAAACTGACGTTGAAAGAGAAAGGAGACAAAGCCAGACAAAGTGAAGTGA 3971
Dh 4194 GAGATGTGATTCAGTTCGAAAGGAGAAACAAAGACAGAGCCCAATGTGTAATGAA 4253
Qy 3972 GAAGG---TGAGCGAGAAACAGCGGCTCTGAGCATGAAGAACCTTACGGGAAGCCAGTC 4028
Dh 4254 GAGAACTTGAAGCAGAAACAGCTGTATCTGATGAAAGGTCAATACAGACTCTC 4313
Qy 4029 CTGACACTTGAACATGACCAGCTCAGAGAGGGGAGAGCACTGGGAAGCTTTGAGAGAGC 4088
Dh 4314 CAGACAGTGAATGTGCTCATATGATGAGGCGCAAGAAAGTCAACAGTGTGAAAGAGAC 4373
Qy 4089 CCTTCTC---TCCAGACCAAGAACAGAGGTTGATAGAGTTCAAGTTCAAGCTG 4145
Dh 4374 CTTCTCTCTGCTTGAAGTCAAGAGAGGAGTATGACCAAAATTCAGATCAAGAGCTCT 4433
Qy 4146 GACACAAAGTCACTCAAAACGACAGAGCTGTGAAAA-----GCTATGAAACGTT 4199
Dh 4434 GAGGCTATCTCACTTAAACGCGGCTGACAGAGGAGAAAGTCTTAAAGGAAATGACC 4493
Qy 4200 GTGATTTCAAGACAGAGTGAAGTCAAGAGTGTGAGTGCACATTTATACAGCTGAG 4259
Dh 4494 AACATTTTGAAGACAGTGAACGTTGAGAGCTGACAGGTGACATTTAGTTCTGGAAGAG 4553
Qy 4260 AAGTCTCTGCAACGGGTGGCACTGACCTTTCAGATGACAGAGACAGCTGACCTGTG 4319
Dh 4554 AAATCTCTGAAAAAATGAAAGACTTGTCCGCTCATCCAGGGGAAAGATGTGTGCCACA 4613
Qy 4320 GGGCTGAGTCTCAGGAGAGATCCATCCCATCATATGATATCTCTGCTCTGAAAGCAC 4379
Dh 4614 GGGCCCGAGCTGCAGGCAAAATTCACACAGTGAATGATGTGCTACTACCAAGAAAGGC 4673
Qy 4380 CTACATCTGACCTTACAGAGAAATTAAGCGATCCCAAGAGAGAGCATCAGAGAGAG 4439
Dh 4674 TTAAGTTCCGACTTGAAGAGAGAGAAACACATCATCAGAAAGTGAAGTCAATGATG 4733
Qy 4440 GACAAAGCAGATGTGCTGTGATGCTGACGGAAGAGAGTACGCAATGCAAAAAAGTC 4499
Dh 4734 GATGAGAGGTTGCTTGGC-----AGAGGTCAAAAGTGAAGTGTGCAATTTG---AGAT 4784
Qy 4500 CTCAAGGCTGAACCTGAGATCTGTGAATCTTGAAGTGAAGTGAAGCAAAATTTCTGTGAC 4559
Dh 4785 TTAGAGCTTGAAGAAATGGAATTTTGAAGCTTGAAGCAAAAGCAAGTAACTTGTCCAAAC 4844

QY 4560 GTCAATCAGACGCGTTGACCGTTCGACGTAC---AGAAACAGCCCGGAACTCAT 4616
Db 4845 ATCATCAGACGCGTTGACCGTTCGACGTACGAGAGAAACAGCCCGGAAATGTTG 4904
QY 4617 GCTTATGATTCACAGACCCAGGTTTC-----CTGCATGCAGGCTTGACAG 4660
Db 4905 ACCTCTGAGTTACAGACACAACTCAGCTGATAAAGCTGACAGCCAGGAGCTGACAG 4964
QY 4661 CAGGAGCCCAACAGATGCTGGACAAAAATGAAAGATGCAAGATGAAACACCCAGTGCC 4720
Db 4965 GAAACGAGAGAAAGAGAGAGAACTCAGGCTCTGCACAGAGTGAACACCAATTAAT 5024
QY 4721 GCAGCCCA-----CAGAGGACTTCCAGTCTGACCGTTCCTGGA 4759
Db 5025 TCAGCCAAAGAGGAGTCAAGTCAACCGCAGTGGGCAAGCAATCTGATATATCCAAA 5084
QY 4760 GGATGGGCTCAGCCTCGGAAATGCTTGGCCGCTTCAGTGTAAAGCGCGGTGTCAA- 4818
Db 5085 GACATGAGTGAAGCTCAGAAAGACCATGCTGTTGAGGTAGAGGTTCCACTGTAAAT 5144
QY 4819 ----- 4818
Db 5145 GATCAGCAGCTGGAAGAGTGTCTCTCCATCTGAGGAAGAGGAGTGGAGCTGGAACA 5204
QY 4819 -----AGTAAGCATTGAGAGCTGCCT 4840
Db 5205 AAGTCTGTGCGCAGAGATGATGTGCTTGTGATGACCAAGAAATAGAGAGTCACTA 5264
QY 4841 CCTCAACCCCAAGA--TCCAAAAGGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTA 4898
Db 5265 GTTGAACCGAAGAGATGAAAGGTGATGTTGATGACCTGAAACCCAGAACTCA 5324
QY 4899 GCCAGGCAGAGGCCAGTCTCTGGAACCTTAACCAAGAAATCCCCAGACACCAAGCGGA 4958
Db 5325 GCCCTGGCTGATCTGATGCTCAGGAGGCTTAACCAAGAGTCCCCAGATACAAATGGA 5384
QY 4959 CCAAGCTAACCGAGAGGCGGATCCCCAAAGTTGAGTGCAGAGAAAGAAATGTCT 5018
Db 5385 CCAAAACAAAAGAGAGAGGAGTGGCCAGGAAATGAGATTGAGGAAAGGAAAGTGCAC 5444
QY 5019 ACCAAGTCAGTCAAGAGAAACAGGCCAGGAGAGAGGAGCTGCAGGAGCCAAAGGGA 5078
Db 5445 AGTGAATCAGATAAGCGATCACCCCAAGCAGGAGAGGTTACAGAAACAGAGAGA 5504
QY 5079 GACTGCGCAGAACTCTAAGATGTTAGTTG-----CTCAATTTGATC 5117
Db 5505 GAATCTGCAAGTCAGAACTTACAGAACTTTAAACATCATGTCAGTTAACTCATTTGCT 5564
QY 5118 ATCTGAAGACCAAGATGTGAAACAAAGTCACAGAACAA-----GATGCTGCTGTGGGAC 5173
Db 5565 GTTTGGAAGACCAAGATGTGAAGACAAAGTAGTAGAGAAATGAATGCTGCTGCTGAGA- 5623
QY 5174 CTTGAGACCAAGATTTTCAGAGCCATGAGATCCAGAGAGCAGGCGCTCCAATGATTTCC 5233
Db 5624 CTGAAGACCAAGATTTTCAGAA-ACCTTGAAGATTGGAGAGCAGGACATCAACTGATCTCA 5682
QY 5234 ACCAGTAGAGCAGCCCGCAATTTCTGAGGCTTCATCGGAGCTAGAGCCAGCTAAACATT 5293
Db 5683 TTTCTAGAGAGC-CCCTGCAATCTGAGGCTTCATCAGAGGCTAGAGCCATTTAACTTT 5741
QY 5294 TCCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGATGCGCTCCGTTATTTCTTAACATT 5353
Db 5742 TCCTCTTTTCAAGACCAACCTACAATTTTCCCTTGTAAAC----- 5781
QY 5354 TCCTCGTTTCAAGACTGCTTTGATTTGCCCCCTTGATGCGCTGCTGATTTTCGGATTTTA 5413
Db 5782 -----CATATAAATTTCTGATTTA 5799
QY 5414 AGGTCTGCTGCTTCAACCTGGAAAC-CAATTTCTGCCATACCTAGTTCCACTTCTCAAACT 5472
Db 5800 AGGTCTTAAATCTTAACCTGGAACTGGAGTTGGCAATACCTAGTTCTGCTTCTGAACT 5859
QY 5473 GGAGCATCTCTTATGATTTATATGATGTTTTATGATGTTTTATGATGATGCTCTCTCTGATACCTAT 5532

Db 5860 GGAGTATCATCTTTTACATATTTATATGATGTTTTAAAGTAG---TCCTCCTGTATCTAT 5916
QY 5533 TGTATATTTTTTTTCTAAACGTTT-----AAGCACATGCTTTTTTGTATTA 5575
Db 5917 TGTATATTTTTTTTCTAAATGTTTAAAGAAATGTGAGGATACTACATGCTTTTTGTATCA 5976
QY 5576 TGCATATATAACGGGTGTGCAGCCATAGCAGCTTTGAAAAGCTCCAAAGCCTCAACTG 5635
Db 5977 CACAGTATATGATGGGCATGTGCCATAGTCAGAGGCTTGGGAGCTTTAAGCCTCAGTTA 6036
QY 5636 TAACCTGCGAGCAACAGATAACATTC-----CTGCAAGAGAGACAGACTCTTTTT 5686
Db 6037 TATAACCCCAAAAAACAGAGCCTCCTAGATGTAAATCTCTGATCAAGGTACAATTTCT 6096
QY 5687 TAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGAAAAGTGGTTGTTTTCT 5746
Db 6097 TAAATTTCTACTAATGATGAGGTCCATATTTAGTGTACTCTGAAAATTTGGTCACTTTCT 6156
QY 5747 ATGCACAGCGAGCTCAGAAATAAAACCCCATTTTGAACATCCAGGATGTCCCAATATT 5806
Db 6157 ATTACACGAGGTGTGCCAAAACATAAAAGCAATTTTGAACATACAGAAATGTTCTATTGTC 6216
QY 5807 ACCATGATTTTTTCCCCCTTTTGTCTAATCCAGTCCAGGTTGGAAAGAGTCTCCTCTG 5866
Db 6217 ATTGGGAAATTTTGC-----TTTCTAAACCCAGTGGAGGTTAGAAAGAAATTTATTTCT 6269
QY 5867 TGTTCAGA-----TTAAGCCCTGCTCTTAATGATATGACAAATAGT 5909
Db 6270 GGTAGCAAAATTAACITTTACATCTTTTCTACTGTTTGGTTGTTGGACCGATAAGT 6329
QY 5910 GTGCCTTAAGGCCCATGAGA-----TGTTCCTTAATGCAGAAAGGAATCTGT 5953
Db 6330 GTGCTTAATCTCTGAGCAAGTAGTGAATATGTTTATATGTTATGAAGAAAGATTTGT 6389
QY 5954 TGTACGTTTTTTTGTGATGTAATCTTCTATGCTGGACCGAATTCATATGAGATGCGAAGTG 6013
Db 6390 TGTAAAG--TTTTGTGATTTCTACTCTTATATGCTGGAATTCACATGCGCATGAAATA 6447
QY 6014 AGTCCTGTTCTTTACAGATGTTATTTGATAGACTGAGTTTGTCTGTTGTTATATCTG 6073
Db 6448 AGTCAGGTTCTTTACAAATGGTATTTTGTATGATGATGATGTTGTTGTTGTTGTTGTTG 6507
QY 6074 TGCCCTCTTTTAAAGAACAAATGTCATTATGTTCTTTGGATAAATTTGATTTGACAA 6133
Db 6508 TGCCATTCCTTTTAAAGAACAAATGTTGCAACATTCATTTGGATAAGTTGTTGATTGACGA 6567
QY 6134 CTGATTTTAAATAAA 6147
Db 6568 CTGATTTTAAATAAA 6581
RESULT 12
AF001504
LOCUS
DEFINITION
AF001504
Homo sapiens myasthenia gravis autoantigen gravin mRNA, partial cds.
ACCESSION
AF001504
VERSION
AF001504.1
GI:2160783
KEYWORDS
Homo sapiens.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 6297)
Bowditch,R.D. and Ginsberg,M.H.
Sequence of gravin cDNA isolated from a human umbilical vein endothelial cell library
JOURNAL
Unpublished
AUTHORS
2 (bases 1 to 6297)
Bowditch,R.D. and Ginsberg,M.H.
Direct Submission
TITLE
Submitted (29-APR-1997) Biochemistry, University of Oklahoma Health

Sciences Center, PO Box 26901, BMSB 853, Oklahoma City, OK 73190,
USA

FEATURES
source

Location/Qualifiers
1..6297
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="endothelial cell"
/class_type="umbilical vein"
1..5226
/note="kinase scaffold protein"
/codon_start=1
/product="myasthenia gravis autoantigen gravin"
/protein_id="AAB58938.1"
/db_xref="gi:2160784"

CDS

translation="DPAIADSPATKLLQKNGSLTINGVADEPDLSEQBDLNGKC
ALNCGALNSQEEVIVTEVGRSDSEVSRDSDKMATSAVAVHITDQSETE
IIEOIPSENLIELTOPTESQANDIGKVFKEKVFYKDKTKSPDVLTVK
KDEBGAAGAGDHDDPSLAGEAASKESEPKSTKEPELTKRQSHALISPPASQ
AVECKEKEKQKESKESKESPSPTSSETSPKPKFTQGNARKKTSFKPK
DEVAISEKKEQEPKVDTEBDGKAENASEKLTASBOAHPOEPASHEPRLSEYK
VELPSEVSGSGSPSEKPAPLATEVFEDEKLEHVEVNAEIVSTVEETBEOKTE
VEETASGVPAELVEMDAEPQEAAPAEVLKLTQVSGEDPTQADLSPEKVLSPK
PEGVASEVEMLSQERMKVQSPKLTFTSTGLKLSGKQKGRGGDESGHTV
PADSPDQEEKGSASSPSEPEITCLEGLAVOODGAEAGATSDGKSEGT
PMASPKMTPTKRVKVRSPSEDEKDELTKVKSATLSTSESTASHEMKGVS
EPRKVDTSVMEALICVSSKKRARRGSSDEGPKAMGSHQKADKDKKGT
TDGLIAGQEHDPQGSPPQAGSPTEGEVSTWESFKRLVTRKSKLEKSESD
SISAGVHSTPDTEPGKEBSWVSKKIPRRKRPRDGKQEPAGEVATGANEED
SDPAVPLSEYDAVERKMEAOQAQKSAEPEQKALAEVSEKLSAEPPVTEPLP
GTRPATLIEERSPSWISASVTEPLEOVAEALTEVLAERLAESEPPVTEPLP
NREARGDVVSBAELTPAATVATAGPILGABEGTGAABETTEMVAVALDSDP
TTEETVQVEVGVPEDEORPTQELQAVAEKESOLPTGSGEDVLOVORA
EAEREEOEASGLKETDVLKVDQAEKTEPTQGVQVQDTBESERKPYTESI
ESSBLVTCOAEETLAGVSOEMVMEQALPPSVETPTDSTGSPVADFPAPGTOK
DEIVEIHEENEAAGTSGGTEAEAVPAKERPAPSPSEFQOEETKESKDELT
DKVEVETVSTLSTKTEGQADQVADETKQVPPFEGLEGIDTGIIVSRKTEVAL
KGEETBAECKDDALEOSHAKSPSPVEREMVQVREKTBAPTHVNEKLEHT
AVTVEESKOLLQTVNPIIDQAEVSSLESGPPCCGOEBAVOTKIQVONSERSPT
LTAABEEKVLGETANILLETETLEPAHALVLEKESVEBOQCAVKSVAIEBEP
COAKSTPIVATITTKGSLDLEKETSLLKMSSEVBEQADVPKSGEDVAPSPG
ENGILEETKSKLVQNIQTAVDQVRETEATLSELTQAHVILKASODAGOE
TEKIEEEOASADDEPTITSAKESSESTAVGOAHEDIKMSSEAKETVTEVESV
NDQILEEVVLSSEEGSGAGCTKSPVEDDGHLLARIKSLIVERPEDEKGDVDPEN
ONSALADTDSAGLTKESPDTNTPKQKEKEDQAEVELDQKHSBDALITPQAEEL
QKQERESAKSELTES"

BASE COUNT 2051 a 1310 c 1708 g 1228 t
ORIGIN

Query Match 34.7%; Score 2136.8; DB 9; Length 6297;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 4120; Conservative 0; Mismatches 1747; Indels 525; Gaps 41;

153 GACCCCGCGAGCGGAGACCCCGCCACCAAGCTCCACAGAAAGATGGCCAGTGTCTT 212
Db GCGATGCTGCTCGGACCCCGCCACCAAGCTCTTACAGAAAGATGTGACGTCTCTCAC 66
7 GCGATGCTGCTCGGACCCCGCCACCAAGCTCTTACAGAAAGATGTGACGTCTCTCAC 66
213 GTCAACGCGTAGCTGAACAAGAGATGTCATGTCACAGAGAAAAACAGAGAGGAC 270
Db ATCAATGGCTAGCTAGCAAGATGAGCTCAGGCTCAGAGAGGTCACCTTAATGGCCAG 126
67 ATCAATGGCTAGCTAGCAAGATGAGCTCAGGCTCAGAGAGGTCACCTTAATGGCCAG 126
271 -----AGAGAGAAAGATGCTTGAAT 290
127 AAAGAGCCCTGAACGCTCAAGAGCCCTTAACAGCCAGAGAGAAAGAAAGATGCTG 186
291 GAGAGATTGGAAGAGAGATGGAAGATGGAAGAAAAAGCCAGTTGAAGAAAT 350
187 ACAGAGTTGGAAGAGAGATGGAAGATGGAAGAAAAAGATGCTCGATTAAGAGATG 246
351 GCGGCCAATCTCAAGCTGTTGAAGATATCAAAAGATGGGACAGAGAGACATCAGAA 410
Db GCTACTAAGTCAGAGGTTTTCACGACATCAGATGATGGGACAGAGACACCCGAA 306
411 ATTAATTGAACAGATCCCTGCTTCAAGAAAACAATGTGAGAAATGTGACAGCTGCTGAG 470

307 ATATCGAACAGATTCTCTTCAGAAAGCAATTTGMAAGCTTAACACACCACCTGAG 366
471 TCCAGGCTATATGATGTTGGCTTCAGAAAGATTTAAATTTGGTGTAAATTACG 530
367 TCCAGGCTATATGATGTTGAATTTAAAGAGTGTAAATTTGGTGTAAATTAC 426
531 GTGAAGAGATTAATAAATGAATAAGTCAATATCTGTCACACTACTGTCAAGAGAT 590
427 GTGAAAAAGATTAAGACAGAGAGACCTGACATCTGACATCTGTAAGAAAGAT 486
591 GAAGGCAAGGGGACAGAGACCTGTCGAGCTGAGACCAAGAGAGCCCACTGTGAG 650
487 GAAGGCAAGGGGACAGAC-----GGGGCTGGCCACACAGGAGCCCAAGCTTGGG 537
551 ACTCCGCTCGAGAGTACAGATCCCAAGAAAGTGAAGTGAAGCAATCCACAGAGAACAA 710
538 GCT-----GGAGAAAGACAGATCCCAAGAAAGGCAACCAATCTTACAGAGAAACC 591
711 GAAGGCAACCTGAAGCAAGAAACAGAGACGACACAGAAATCCCTTCAAGCCGATCTGAT 770
592 GAAGGCAACCTGAAGCGTGAAGCAAGCAAGCAAGAAATTTCTCCCGCAGATCTGGC 651
771 CAAGCGCTGAAGAAAGCCAAAGATGAAGAGAAAGAAACAAAGAAAGACCCAC 830
652 CAA---GCAGTGAAGGAATGCAAGAGAAAGAGAGAAAGAAACAAAGAAACCTAGC 708
831 AAGTCCCAAGATCCCGAGAGACCCCACTCAACGTGAGACAAATCTTCTTCAAGAG 890
709 AAGTCTGAGAGATCTCCGACATGATCCGCTGACAGTGAAGAGATCAACTTCAAAAAA 768
891 TTCTTCACTCAGGTTGGGCGCGCTGGCGCAAGAAACCACTTCAAGAAATCAAAAGAG 950
769 TTCTTCACTCAGGTTGGGCGCGCTGGCGCAAGAAACCACTTCAAGAAACCAAGAG 828
951 GATGATCTGAAATCTCCGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1010
829 GATGAAGTGAAGCTTCAGAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 888
1011 GAAAAAGAAAGACAGAGCCAGCTCCGAGAG-----G 1043
889 GAAGGCGAAAGGACAGAGAGTTCCTCCAGAAATCTGACCGCTCCGAGAAACCCACCA 948
1044 CAGAGCGCGGACAGAAACAGACAGACAGAGCCAGGTTGTGACAGACATAGAGAAAGTGAG 1103
949 CAGAGCGCGGACAGAAAGTCCACAGAGCCCGGTTATAGCTGAATAGAGAAAGTGAG 1008
1104 CTGCTTTGAAAGACAGAGTTGTGACCTGAGAGGATGTCAGAGAGAGAGAGAGAG 1163
1009 CTGCTTCAAGAGAGACAGAGTCAAGTGGCTGCGAGGAGACCTTCTGAAGAGAAACCTGCTCG 1068
1164 TTGGCAAGAGAGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1220
1069 TTGGGCAAGAGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
1221 GTCCACGTGAGACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280
1129 GTCCACGTGAGACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1170
1281 GGGGCGTGTGTGAAGAGAGAGAGAGAGAGAGATCTTGGCCCTGAGAACTGGCTGAGCCC 1340
1171 -----AGGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1224
1341 CAGAGAGTCCCGAG 1400
1225 GATGCAAGAACTCAG 1284
1401 GTCTCTGAG 1460
1285 GTTTCGAG 1344
1461 AAACACCCAGAGAGAGATGTCAGTGAAGTGAAGTGTCTCTCAAGAGAAATCAAG 1520

Db 1345 AACCCCCGAAGCGTTGTGAGTGAGGTGGAATGCTGTCTATCACAGGAGAAATGAAG 1404
Qy 1521 GTACAGGAAGTCCCTTGAAGAAACTCTTTCAGTAGCTCAGGCTTAAAGAAAGCTCTCTGGG 1580
Db 1405 GTGACGGGAAGTCCACTAAAGAACTTTTACCAGCACTGGCTTTAAAAAAGCTTCTTGA 1464
Qy 1581 AAGAAGCAGAGGGAAGCAAGAGAGGTGGGGAGACCAAGAGCCTGGAGATACCAACAC 1640
Db 1465 AAGAAACAGAAAGGAAGAGGA--GGAGGACACAGGAATCAGGGGAGCACACTCAG 1521
Qy 1641 ATTACACCCGAATCCCGAGAGAGTCTGATGAGCAGAAAGGAGAGAGCTCTGCGTCCGTCC 1700
Db 1522 GTTCCAGCCGATTCTCCGACAGCCAGAGGAGCAAAAGGGGAGAGCTCTGCTCTCATCC 1581
Qy 1701 CCGAGAGCTCTGAGAGACCACTGTCTGTGAGAAAGGCGCTGGAGAACCCCAAGAT 1760
Db 1582 CCTGAGGAGCCGAGGAGATCACGTGTCTGAAAGGGCTTAGCCGAGGTGCAGCAGAT 1641
Qy 1761 GGGGAAGCTGAGGAAGAACTACTTCCGATGGAGAGAGAGAGAGAGGATCACTCCC 1820
Db 1642 GGGGAAGCTGAAGAGGAGCTACTTCCGATGGAGAGAAAGAGAGAGGTGTCACTCCC 1701
Qy 1821 TGGGCATCTTTCAAAAGATGGTGACACCCCAAGAAACGGGTCCGAAGACCTTCTGAGAT 1880
Db 1702 TGGGCTTCATTCAAAAGATGGTGACCCCAAGAGCGGTGTAGACGGCTTCGAAAGT 1761
Qy 1881 GACAGAGGAAGAGCTGGAGAGGTCAAGAGCCCACTTGTCTCTCACTGATAGACACA 1940
Db 1762 GATAAAGAAGATGAGCTGGACAGGTCAAGAGCTACCTTGTCTTCCACCGAGAGACA 1821
Qy 1941 GTGTCAAAATGCAAGATGAAGTCAAACTGTTGGTGAGGAAACAAAGCCAGAGAAACA 2000
Db 1822 GCCTCTGAAATGCAAGAGAAATGAAGAGGAGCGGTGGAAGAGCCAAAGCCGGAAGACA 1881
Qy 2001 AAGCTAGGTTGGATCTTCACTGTCTTGGGAAGCACTGATTTGTGTGCGATCATCCAAG 2060
Db 1882 AAGCGCAAGTTGATCTCAGTATCTTGGGAAGCTTAAATTTGTGTGGATCATCCAG 1941
Qy 2061 AAGAGCAAGGAAGGATCTCTTTCAGATGATGAAGAGGGGCCAAGGACACTGGGAGG 2120
Db 1942 AAAAGACAAAGGAGGCTCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGGAGGA 2001
Qy 2121 GACAGTACAGAGCAGAGGAGGCCAGCAAAAGACAAAGCCGGAACAGAGCTGTTCCT 2180
Db 2002 GACCACAGAAAGCTGATGAGGCCGGAAGAACAAAGAGACGGGACAGACGGGATCCTT 2061
Qy 2181 GCCAGCACCCAGGAGCAGGACCAAGCCCAAGGAAGTTCTCACCCGAGCCAGCGGGAAGC 2240
Db 2062 GCTGTTCCCAAGAACATGATCCAGGGCAGGAAAGTTCTCCCGGAGCAAGCTGGHAGC 2121
Qy 2241 CTTTCCGAAGGGGAAGGTGTCTCCACTTGGGAGTCAITTTAAAGATTAGTCACTCCAAGA 2300
Db 2122 CTACCCGAAGGGGAGGCGTTTCCACCTGGGAGTCAITTTAAAGGTTAGTCACGCCAAGA 2181
Qy 2301 AAAAATCCAAAGTCAAACTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTA 2348
Db 2182 AAAAATCAAAAGTCCAAAGCTGGAGAGAGAAAGCGGAAGCTCCATAGCTGGGTCTGGTGTA 2241
Qy 2349 GAGCAGTTGTCCACTGAGATCGAACCCAGTAGAGAAATCTTTGGGTTTCCATTAAAGAAA 2408
Db 2242 GAACATTCCTCACTCAGACACTGAACCCGGTAAAGAGAAATCTGGGTCTCAATCAAGAG 2301
Qy 2409 TTCAATCCCGGACGGCGGAAGAAAGGGCAGACGGGGAAGCAAGAACCAAGCACTGTGGAA 2468
Db 2302 TTTATTCTTGGACGAAGGAAGAAAGCCAGATGGGAAAACAGAACAAAGCCCTGTGTAA 2361
Qy 2469 GACTCAGGCGCAGTGGAGATAAATGAGGACGACCTTAATGTCCAGCGCTGTGCTCTG 2528
Db 2362 GACGACGGGCAACAGGGGCCAACGAAGATGACTCTCATGTCCCGGCGGTGGTCCCTCTG 2421
Qy 2529 TCTGAGTATATGACAGTGGAGGGGAGAGATGG-----AAGCCAGGGGAATACG 2579
Db 2422 TCTGAGTATGATGCTGTAGAAAGGGGAGAAATGGAGGCAAGCAAGGCCCAAAAGCGCA 2481

Qy 2580 GAGCTGCCCCAGCTGTCTGGGGGCTGTGTACGTGCCAGAGAGCTCAGTAAGACTCTGGTC 2639
Db 2482 GAGCAGCCCGAGCAGAAAGGCGAGCCACTGAGGTGTCCAAAGAGCTCAGCGAGAGTCAGGTT 2541
Qy 2640 CACACTGTGAGTGTCCAGCTCATTGATGGACCCAGGGCAGTCAACAGTGTCCGAAGCGG 2699
Db 2542 CATATGATGGCAGCAGCTGTGCTGAGCGGACGAGGGCAGCTACCATTTATTGAAGAAGG 2601
Qy 2700 TCTCTCTTCTGGATATCCGCTTCCGTTAAACAGAACTCTTGAACACACAGCGGAGAGGCC 2759
Db 2602 TCTCTCTTCTGGATATCTGCTTCAGTGCAGAACCTCTTGAACAAGTAGAGCTGAAGCC 2661
Qy 2760 ATGCCACCTGTTTGAAGAGTCACTGAAAGAGACATCATTTGC---AGAAGAAAATCTCTGTG 2816
Db 2662 GCACCTGTAACTGAGGAGGTATTGGAAAGAGAAATTAATTGCAGAAAGAAACCCCCACG 2721
Qy 2817 CTACCCAGACGTTTACAGAGGTTAAAGATGCCCATGACGACATGTCACCACTGAAGTG 2876
Db 2722 GTTACTGAACCTTGCACAGAAACAGAGAGGCCCGGGGCGACACGGTCTGTTAGTAGGGC 2781
Qy 2877 GATTTCACTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAA 2936
Db 2782 GAATTCACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGGCCATTGGTGCAGAGAA 2841
Qy 2937 GTTACCGAAGCATCTGGGGCCGAAGAGACACAGACATGTTGTCGACGTTTCCCAAGTG 2996
Db 2842 GGAACCGAAGCATCTGCTGTAAGAGACCAACAGAAATGTTGTGCAGCAGTCTCCCAGTTA 2901
Qy 2997 ACTGACTCCCCAGACACCAACAGAGGAAGCCACCCAGTTTCAGAGGTAGAGAGTGTGTG 3056
Db 2902 ACCGACTCCCCAGACACCAAGAGGAGGCCACTCCCGTGCAGGAGGTGGAAGGTGGCGTA 2961
Qy 3057 CTAGATACAGAAAGAGAGGAGCCAGAGCGAGGCCATCTCCAAGCCGTTCGAGACAAG 3116
Db 2962 CTTGACATAGAGAGCAGAGAGGCGGACTCAAGAGGTTCTCCAGCAGTGGCAGAAAAA 3021
Qy 3117 GTGAAAGAGAGTCCAGGTGCGCTGCAACC---CAGACTGTGCAGAGAACGGGGTCAAAA 3173
Db 3022 GTGAAAGAGGAAATCCCAAGTGCCTGGCACCGGTGGCCAGAAAGTGTGCTTCAAGCTGTG 3081
Qy 3174 GCACCTGGAGAGTTGAGGAGGTAGAGGAGACTCCGAAGTGTGCTTCCGGAAGAAAGAG 3233
Db 3082 CAGAGAGCAGAGCAGAAAGACCCAGAAAGAGCAGGCTGAAACGCTCGGGTCTGAAAGAAAGAG 3141
Qy 3234 AAGGACGTTATGCCAAAGACCCGTGCAGGAAGCTGGAGCTGAGCATCTTGCACAGGGC 3293
Db 3142 ACGGATGTAGTTTGAAGTAGATGCTCAGGAGGCAAAAACCTGAGCCTTTTACAAAGGG 3201
Qy 3294 TCTGAGACTGGACAGGCTACTCCAGAGAGCCTTG---AAGTTCTCTGAAGTCACGGCAGAT 3350
Db 3202 AAGGTGTTGGGCGAGACACCCCGAGAAAGCTTTGAAAAAGTCTCTCAAGTCACAGAGAGC 3261
Qy 3351 GTAGAC-----CATGTGCGCCAGTGCCA-----GGTTATCAAG 3383
Db 3262 ATAGAGTCCAGTGAGCTTTGTAACCACTTGTCAAGCCGAAACCTTAGCTGGGGTAAAAATCA 3321
Qy 3384 CTCACAGCTGATGGAACAGGCGGTGGCCCTGAGTCAATCGAAACCTTGCAGACAGT 3443
Db 3322 CAGGAGATGTTGATGGAACAGGCTATCCCCCTGACTCCGTTGGAACCCCTTACAGACAGT 3381
Qy 3444 GAGACAAATGGAGACACTCCCTTAGCAGATTTCAGACATCTGCAGATGGGACACAGCAAGAT 3503
Db 3382 GAGACTGATGGAAAGCACCCCGTAGCCGACTTTGACGCACCGGACACACCCAGAAAGAC 3441
Qy 3504 GAAACCATTTGACAGCCAGGACAGTAAGCCACTGCAGTGTGCAGGAGTCAAGGTCACA 3563
Db 3442 GAGATTGTGGAATAATCCATGAGGAGATGAGTGCATCTGGTACCCAGTCAAGGGGCA 3501
Qy 3564 GAAGAGAGCGGCTACTGCTCAGAAAGAGGAGCCTTCACACTACTTAATATGTTCCA 3623
Db 3502 GAAGCAGAGGCGAGTTCTCTGCACAGAAAGAGAGCGCCTCCAGCACCTTCCAGTTTGTGTTTC 3561

```

OY 3624 GCCCAGAGAACATGCGGAGAAACGAGAAAGATGTTCTTGAACCTTACACAGCAAG 3683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3562 CAGGAGAAATTAAGAAACAATCAAAAGATGAAAGACCTTAGAGATACAGATTAAGAG 3621
OY 3684 CTTACGTGAGCGCGTCCGTTCTGSCAAAGACTGAGTGGGTCAAGAGGTGAGGTT 3743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3622 GTGTAGTGAAGAACTGTATCCATTTCTGCAAGACTGAGGGGACTCAAGAGGCTGACAG 3681
OY 3744 GACTGTGTGATGAGAAAAAGTCAAA-----GAAAGACAGAGGTGTTGTACAC 3794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3682 TATGTGATGAGAAAAACCAAGAGTACCATTTTTCGAAGGACTTGAAGGGGTCTATAGAC 3741
OY 3795 TCTGG-----ACCCAACATCAAAAGGCTGTGATGACATATGACAGTGAAGTATG 3848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3742 ACAGGCATACAGTCACTCGGAGAAAGTCACTGAAGTTGCTTAAAGGTGAAGGACACA 3801
OY 3849 GGAGTGGCGCGGTGTCAAGAAAAAGAGTACTGAAGTGCAGAGTCTTAG----- 3898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3802 GAAGAACTGAATGTAAAAAGATATATGCTCTTGAATCTGCAAGTCAAGCTTAAGTCTCT 3861
OY 3899 -----CTGAGAGAGAGAGATGAGAACTGACGTTGAAAAAGAGAAAAAGGAGACAAAG 3953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3862 CCATCCCCCGGTGAGAGAGAGATGATGTTCAAGTCAAGAGGAGAAAAACAGAACAGAG 3921
OY 3954 CCAGAGCAAGTGAAG--AAGGTAGCAGAAACAGCCGCTCTGAGCATGAAGGA 4010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3922 CCAACCCATGTGAATGAAGAGAGAGCTTGAGCAGAAACAGCTGTATCCGTATCTAAGAG 3981
OY 4011 ACCTAGGGAAGCCAGTCTGACCTTGACATGCCAGCTCAGAGGGGAGGAGCACTG 4070
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3982 GTCAATAGCACTCTCTCAGACAGTGAATGTGCCATATATGATGGGCAAAAGAGTGC 4041
OY 4071 GGAACCTTGAAGAGAGCCCTTCTCT---CCAGACCAAGACAAAGCAGTGTCAATAG 4127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4042 AGCAGTTTGAAGAAAGCCCTCTCTCTGCTTAGTCAAGAGAGAGATATGACCAAA 4101
OY 4128 GTTCAAGTTCAAAAGCTGAGCAACAAGTCACTCAAAACGACGAAAGCTGTGAAAA-- 4183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4102 ATTCAAGTTCAAGGCTGAGGATCATCTTAACAGCGGCTGCAAGAGAGGAAAAAG 4161
OY 4184 --GCTCATGAAACGGTGTGATTCAGAGACAGGTGAAGTCCAGAGTGTGATGTCGA 4241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4162 GTCTTAGGAAATGCGCAACATTTTGAACAGGTGAACGTTGAGGCTCAGAGTGA 4221
OY 4242 CACTTATTAACAGTGAAGTCTCTGCAACGGGTGGCCACTGGACTTTGACATGCA 4301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4222 CATTTAGTTCTGGAAGAGAAATCTCTGAAAAAATGAAGACTTTGCGCTCATCAGAG 4281
OY 4302 GAGGACCGGTACCCCTGGGGCTGAGTCTCAGGCGAATTCATCCCATTCATTAAGT 4361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4282 GAAAGTGTGTGCCCAACAGGGGCCGACTGTCAAGCAAAATCGACACAGTGAATATCT 4341
OY 4362 CCTGCTCCGGAAGACCCCTACATCTGACCTAACAGAGAAATAAGCCATCCAGAGA 4421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4342 GCTACTACCAAGAAAGGCTTAAAGTTCGACCTGGAAGGAGAGAAACCAATCATCTGA 4401
OY 4422 GAGCGATCAGAGAAAGAGACAGCCAGATGCTGATCTGATGCTGACGCAAGAGAT 4481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4402 TGGAAATCAGATGAATGATGATGAGAGGTGCTTGCC-----AGAGAGTCAAAAGTGA 4455
OY 4482 ACAGCAATGAAAAAGTCTCTCAAGGCTGAACCTGAGATCTTGGAATTTGAGATAGAGC 4541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4456 GTAGCAATGAGAGGATTTAGAGCTGAAAAATGGGATTTTGGACCTTGAGACCAAAAGC 4515
OY 4542 AACAGATTTGTGTGAACGTCATTCAGACAGCCGTTGACAGTTCGACGTAAC---AGAA 4598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4516 AGTAAACTTGTCCAAACATCATCCAGACAGCCGTTGACCAAGTTTATGATACAGAAAG 4575
OY 4599 ACAGCCCGGAAACTCATGCTTATGATTCACAGACCCAGGTTG-----C 4642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4576 ACAGCCACGAAATGTTGAGCTGTGAGTTACAGACACAAAGCTCAGGTGATTAAGCTGAG 4635
OY 4643 TGCATGACAGCTTGACAGCAGGAGCCCAACAGATGCTGACAAAAATGAAGATGCCAA 4702

```

```

Db 4636 AGCCAGACGCTGAGACAGAAACGAGAAAGAGAGAGAACCTCAGGCTCTGACAG 4695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4703 GATGAACACCCAGTCCGACGCCA-----GAGAGACTTGCA 4741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4696 GATGAACACCAATTAATCTCAGCCAAAGAGAGTCAAGTCAACCGCAGTGGACAGCA 4755
OY 4742 AGTCTGACCGTTCGAGGACATGGGCTCAGCCTCGGAAATGTGTGCGGCTTGACGTT 4801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4756 CATTCATTAATTTCCAAAGACATGAGTGAAGCTTCAGAAAGACATGACTGTTAGGTA 4815
OY 4802 GAAAGCCCGGTGTCAA----- 4818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4816 GAAAGTTCACTGAATATGATCAGCAGCTGGAAGAGGTGCTCCCATCTGAGAAAG 4875
OY 4819 -----AGTA 4822
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4876 GAGGTGAGCTGGAACAAAGTCTGTCCAGAAAGTATGTCATGCTTTGACAGAA 4935
OY 4823 AGCATTGAAGAGCTGCTCCCAACCCAAAGA--TCCAAAGAGCATGCTGTGATGGC 4880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4936 AGAATGAGAAAGTCACTAGTTGAACCGAAAGAGATGAAGAAAGGTGATGTTGATGAC 4995
OY 4881 CCTCAGCTCCAAAGCTTAAGCCCAGGAGAGGCGCAGTCTCTGAGAACTTAACCAAAGAA 4940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4996 CTTGAAAAACGAAGTCAAGCCCTGGCTGATCTGATCTCCTCAGAGAGGCTTAACCAAAGAG 5055
OY 4941 TCCCAAGACACCAACGACCAAGCTTAACGAGAGGCGATCCCCCAAAAGTTGAGTGC 5000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5056 TCCCAAGTACAAATGAGCCAAACAAAGAGAGAGATGCTCCAGAGATGGAATTG 5115
OY 5001 CAGGAAGAAAGATGTCTACCAAGTCAAGTCAAGAAAGAAAGGCCAGGAGAGAGAGAC 5060
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5116 CAGGAAGAAAGAGTCAAGTGAATCAGATTAACGATCACACCCAAAGCAGAGAGAG 5175
OY 5061 CTGACAGAGCCAAAGGAGAGCTGAGAAATCTTAAGTGTAGTTG----- 5107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5176 TTACGAAGAACAGAGAGAGATCTGCAAGTCAAGACTTAAGAAATCATCATG 5235
OY 5108 -----CTCATTTATCATCTGTGAAGCCAGATGTGAAGAAACAAATCAAGCAAG-- 5157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5236 CAGTTAACTCATTTGTCTGTTTGAAGCCAGAAATGTGAAGACAAAGTGAAGAAAT 5295
OY 5158 --AAGCTGCTGTGGACCTTGAGACCAAGATTCAGAGCCATGAGATCCAGAGAGAG 5215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5296 GAATGTCTGCTGAGA--CTAAGACAGTATTTTCAAG--ACTTGAGAAATGGAGAGAG 5353
OY 5216 GGCCGTCAATGATTTCCACCAGTGAAGCAACCCGACAAATTCGAGGCTTCATCGGAG 5275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5354 GCACATCAACTGATCTCATTTCTAGAGAGC--CCGTGACAAATCTGAGGCTTCATCAGAG 5412
OY 5276 CTAGAGCAGCTAACATTTCTCGTTTCAAGCTGCTTGAATTTGCCCTTGATGCCGT 5335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5413 CTAGAGCATTTAATTTCTCTTCCMAAGACCACTTCAATTTTCCCTGTAATAC-- 5470
OY 5336 CCGTATATTTCTAATTTCTCTCGTTTCAAGAGTCCCTTGAATTTGCCCTTGATGCCGT 5395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5471 ----- 5470
OY 5396 CCGTATTTTGGATTTAAGTCTGCGCTTCAACCTGAAC--CAATTTCTGCCATACCT 5454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5471 CATATAAATTTGATTTAAGGTCCTAAATTTCTTAACCTGGAACGTGAGATTTGGCAATACCT 5530
OY 5455 AGTTCACCTTCTCAACCTGAGACATCTCTTATGATTTATATGATGATTTATGATG 5514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5531 AATTCTGCTTGAACCTGAGATCATTTCTTACATATTATATGATGATTTAAGTAG 5590
OY 5515 TCCCTCTCTGTAACCTATTTATTTATTTTCTTAACGTTT-----AAG 5557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5591 ---TCCCTCTGATCATTTATGATTTTCTTATATGATTTAAGAAATGTCAGATATC 5647
OY 5558 CACATGCTTTGATTTATGCAATATATTAACGGGTGTGACGACCATAGGACGCTTTGAAA 5617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Qy	1122	GTTCGTGACCTGGAGGGATCTGTCAAGAGGAAAGTGTGCTCCCTTTGGCAACGGAAAGTGT	118
Db	1341	GTCACTGGCTCGCAGGGACCTTCTTGAAGAAAACCTGTCTGCTTGGCAAGAAAGTGT	1400
Qy	1182	GATGAGAAAGATGGAAAGCCCAACGAAGAA---GTTGTTGCAGAGGTCACGTGAGCAACCGTG	1238
Db	1401	GATGAGAAATTAAGATGCCACCAAGAAAGAGTTGTGGCCGAAGTCCACGTCAAGCAACCGTG	1460
Qy	1239	GAGAAAGCAAGAGAGAGCAAGGGAGAGAGAGAGAGCTGAAGGGGGGCTGTGTAGAA	128
Db	1461	GAGAGAGGAACCGAAGAGCAAGAA-----ACGAGAGGTGAA	1496
Qy	1299	GGAAACAGAGAAATCTTGCCCCCTGAGAAACTTGGCTGAGACCCCAAGAGGTCCCCAGAA	1358
Db	1497	GAACACAGAGGCTCTGTGTCAGCTGAAGAAATTTGTTGGAATGGAATGCAAGAACTCAGAA	1556
Qy	1359	GCTGAGCTCTGTGAGAGCTGATGAAGCAGAGAGATGTGTCTCTTGAGAGAGACAC	1418
Db	1557	GCCAAACCTGCAAGAGAGCTGTGTAAGCTCAAGAAACGTGTGTTCCGAGAGAGACCT	1616
Qy	1419	ACTGAACAGACAGACCTAACTCCGAAAGAAAGAGAGCTGCCAAACACCCAGAAAGCACT	1478
Db	1617	ACACAGGAGAGTGAACCTCAGCTCTGATGAGAGGTGCTCTCAAAACCCCCGAAAGCTT	1676
Qy	1479	GTCAGTGAAGTGAAGATGCTGTCTCTCAAGAAAGAAATCAAGATCAGAGAAAGTCCCTG	1538
Db	1677	GTGAGTGAAGTGAAGAAATGCTGTATCAACAGAGAGAAATGAAGGTCAAGGAAAGTCACTA	1736
Qy	1539	AAGAAACTCTTCAGTACTCAGGCTTTAAAGAGTGTCTGGAGAAAGACAGAGAGGGAA	1598
Db	1737	AAGAGAGCTTTTACACACACTGGCTTTAAAAAGCTTTCTGAAAGAAACAGAAAGGGAA	1796
Qy	1599	CGAAGAGGTGGGGAGACGAAGAGCCTGAGATACCAACACATTTCAACCGAATCCCCA	1658
Db	1797	AGAGGA---GGAGGAGACAGAGATCAGGGAGACACATCAAGTTCCAGCCGATTTCTCG	1853
Qy	1659	GAGAGTCTGATGACAGAAAGGAGAGAGCTTGTGCTGCTCCCCGAGAGAGCTTGAAGAG	1718
Db	1854	GACAGCAGAGGGAGGCAAAAGGGCGAGAGCTGTGCTCACTCCCTGAGAGAGCCCAAGAG	1913
Qy	1719	ACCACTGTCTGAGAAAGGGCCGCTGAAAGCAACCCAGAGTGGGGAAAGCTTGAAGAA	1778
Db	1914	ATCACGTGTCTGGAAGAAAGGGCTTAACCCAGAGTGCACAGAGATGGGGAACTGAAGAGAA	1973
Qy	1779	ACTACTTCCGATGAGAGAAAGAGAGAGGAGTCACTCCCTGGGCACTCCTTCAAAAG	1838
Db	1974	GCTACTTCCGATGAGAGAAAGAAAGAGAGTGTCACTCCCTGGGCACTCATTCAAAAG	2033
Qy	1839	ATGTGTACACCCAGAAACGGGTCCCAAGACTTTCTGAGATGAGCAAGAGAGAAAGCTG	1898
Db	2034	ATGTGTAGGCCCAAGAGCGGTGTGAGCGGCTTCCGAAAGATGAATGAAGATAGCTG	2093
Qy	1899	GAGAAAGTCAAGACGCGCACCTTGTCTCTCACTGATAGCACAGTGTCAAGAAATGCAGAT	1958
Db	2094	GACAAAGTCAAGACGCTCACTCTGTCTTCCACCGAGAGCACAGCTGTGAATTCAGAA	2153
Qy	1959	GAGTCAAAACTGTTGGTGAAGAACAAAGCCAGAGAAACCAAAGCTGTAGGTGATATCT	2018
Db	2154	GAATATGAAGGGAGCGTGTGAAGAGCCAAAGCCCGAAGAAACCAAAGCGCAAGGTGATATCC	2213
Qy	2019	TCAGAGTCTTGGGAAGCAGATTGTGTGGATTCATCCAAAGAGAGAGCAGAGAAAGCA	2078
Db	2214	TCAGTATCTTGGAAAGCTTTAATTTGTGTGGATTCATCCAAAGAAAGAGCAAGAGAAAG	2273
Qy	2079	TCTCTTTCAGATGATGAAGAGGGCCAAAGACACTGGAGGGGACAGTCAAGACAGAG	2138
Db	2274	TCTCTTTCAGATGAGAAAGGGGGACCAAAGCAATGGAGAGAGACCAACAGAAACTGAT	2333
Qy	2139	GAGGCCAGCAAAAGCAAAAGAACCCGGAACAGACGCTGTTCTGTCCAGCAACCCAGAGAGAG	2198
Db	2334	GAGGCCGGAAGAAAGCAAAAGAACCGGGGACAGACGGAGTCTGTCTGTGTTCTCCAAAGCACT	2393

OY	2139	GACAAAGCGCAAGAAAGTTCTCTCACCCGAGCAGCGGGAAAGCCCTTCGAAAGGGGAAGGT	2258
Db	2394	GATCCAGGCGCAGGGAAAGTTCTCTCCCGAGCAAGCTGAAGCCCTTAACGAAGGGAGGGCC	2453
OY	2259	GTCCTCACTTGGGAGTCATTTAAAGATTAGTCACTCCAAAGAAAAATCCAAAGTCAAAA	2318
Db	2454	GTTTCCACTTGGGAGTCATTTAAAGTTAGTCAGCCCAAGAAAAAATCAAAAGTCCAAAG	2513
OY	2319	CTGGAAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAAGTTGTCACTGAG	2366
Db	2514	CTGGAAAGAAAAAGCGAAGACTCCATAGCTGGGTCGTGGTGTGAACATTTCCACTCCAGAC	2573
OY	2367	ATCCAAACCGAGTAGAAGAAATCTTGGGTTTTCATTTAAAGAAATTCATCCCGGCAAGCGG	2426
Db	2574	ACTGAACCCCGGTAAAGAAATCTCGGGTCTCAATCAAGAAAGTTATTTCTTGAAGAAAG	2633
OY	2427	AAGAAAGGGCCGACGCGGAGCAAGCAAAAGCACTGTGAAAGACTCAAGGGCCAGTGGAG	2486
Db	2634	AAGAAAGGCCGAGATGGGAACAGAAACAAAGCCCTGTGTAAGACGCAAGGCCAACAGGG	2693
OY	2487	ATAAATGAGAGCAGCCCTAATGTCTCCAGCCGTCGTGCTCTGTAGTATATATGCAAGTG	2546
Db	2694	GCCAAACGAAGATGACTGTGATGTCTCCGGCGGTGTCCTCTGTGTGAGTATGATGCTGTA	2753
OY	2547	GAGAGGGAGAAAGATGG-----AAGCCCAAGGGGAATACGAGACTGCCACGCTGCTG	2597
Db	2754	GAAAGGGAGAAATATGAAAGCACAAGCAACCCCAAAAAGGCGAGACAGCCCAAGCAAGAG	2813
OY	2598	GGGGCTGTGTACGTGTCCGAGAGGCTCAGTAAAGACTGTGTCCACACTGTGAGTGTGCGCA	2657
Db	2814	GCAAGCACTGAGGTGTCCAAAGAGCTTCAGACGAGTCAAGTTCAATGATGGCAGCAAGCT	2873
OY	2658	GTCAATTGATGGGACCAGGGCACTGCACAGTGTGAAAGACGGTCTCTTGTGATATATCC	2717
Db	2874	GTCCGTGAACGGGACAGGGCAGCTACATTAATTAAGAAAGGTCTCTCTTGTGATATCT	2933
OY	2718	GCTTCCCTTAAACAAACCTTTGAACACACAGCGGAGAAAGCCATGCACCTGTGAAGAG	2777
Db	2934	GCTTCAAGTACAGAAACCTCTTGAAACAAAGTAAAGGTGAAGCCGCACTGTAACTGAGGAG	2993
OY	2778	GTCACTGAAAAAGACATCATTTGC---AGAAAGAACTCCGTGTCTCACCCAGCGTTACCA	2834
Db	2994	GTAATTGAAAGAAAGTAAATTTCAGAAAGAAAGCCCCACGGTTATCTGAACCTCTGCACA	3053
OY	2835	GAGGCTAAAGATGGCCCATATGACGACATGTGTCAACAGTGAAGTGAATTTCACTCAGAAAGCT	2894
Db	3054	GAGAAACAGAGAGGCCCCGGGCGCACACGGTCTTAATGAGGGCGGAATTGACCCCCGAAGCT	3113
OY	2895	GTGACACCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAAATTACCGAAGCATGTGGGG	2954
Db	3114	GTGACACGCTGCAGAAATCTGCAGGGCCATTTGGGTTTCCGAAGAAAGAAACCAACATCTGCT	3173
OY	2955	GCCGAAGAGACCAACAGACATGTGTGCCGAGTTTCCACGCTACAGACTCCCCAGACACC	3014
Db	3174	GCTGAAGAGACCAACAGAAATGTGTGTGACAGTCTCCCAAGTTAAACGACTTCCCAAGCACCC	3233
OY	3015	ACAGAGAAAGCCACCCCAAGTTTCAGAGGTGAGAGTGTGTCTAGTATCAGAAAGAAAG	3074
Db	3234	AACAGAGAGGCCACTCTCGGTGACAGAGGTGGAAGGTGGGTACCTGACATAGAAAGCA	3293
OY	3075	GAGCGCAGACGCAAGGCATCTCTCCAGCCGTTTGACAGACAGGTGAAGAGAGTCTCCAG	3134
Db	3294	GAGAGCGGACTCAAGAGGTCTCTCCAGCGAGTGGCAGAAAAAGTGAAGAGAGATCTCCAG	3353
OY	3135	GTGCGCTGCACAC---CAGACTGTGCAGAAACGGGCTCAAAAGCACTGAGAAAGTGTGAG	3191
Db	3354	CTGCTGTGCAACCGGTGGGCGCAAGATGTGTCTTCAAGCTGTGCAGAGACACAGGCGAGAA	3413
OY	3192	GAGGTGAGAGAGACACTCCCAAGTGTGCTTTCGAGAAAGAGAGACCTTATGCGGAAA	3251
Db	3414	AACACCAAGAAAGACAGGCTGAACGCTGTGGTCTGTGAAGAAAGACGAGATGTATGTGTTGAAA	3473
OY	3252	GGACCCGTGACAGAAAGCTGAGCTGAGCATTTTGACAAGGCGCTCTGAGACTGACAGGCT	3311

Db 3474 GTAGATGCTCAGGAGGCAAAACCTGAGCGCTTTTACAAAGGGAAGGTGGTGGGCACACC 3533
Qy 3312 ACTCCAGAGAGCCCTTG---AAGTTTCCTGAAGTCAAGGAGAGATGTAGAC-----CAT 3359
Db 3534 ACCCCAGAAAGCTTTGAAAGAGCTCCTCAAGTACAGAGAGCATAGAGTCCAGTGAGCTT 3593
Qy 3360 GTCCGCAAGTGCCA-----GTTTATCAAGCTTCCAGAGCTGATGGAA 3401
Db 3594 GTAAACCACTTTGTCAAGCGCAAAACCTTAGCTGGGGTGAATAATCACAGGAGATGGTATGGAA 3653
Qy 3402 CAGGCGGTGGCCCTGAGTCAATCCGAACCTTGACAGACAGTGAAGAAATGGGAACACT 3461
Db 3654 CAGGCTATCCCCCTGACTCGGTGGAAACCTTACAGACAGTGAAGTCAAGTGGGAAGCACC 3713
Qy 3462 CCCTTAGCAGATTACAGACACTGCGAGATGGGACACAGCAAGATGAACCATTTGAAGCCAG 3521
Db 3714 CCGGTAGCCGACTTTGACGCCACAGGCACAAACCCAGAAAGACGAGATTGTGGAATCCAT 3773
Qy 3522 GACAGTAAAGCCACTGAGCTGTGAGGAGTCAAGGTCAAGAAAGAGGCGGCTACT 3581
Db 3774 GAGGAGAAATGA---GGTGCATCTCGTATCCAGTCAAGGGGCACAGAAAGCAGAGGCGAGTTCCCT 3830
Qy 3582 GCTCAGAAAGAGGAGCCTTCGACACTACCTAAATAATGTTCCAGCCAGGAGAACATGGG 3641
Db 3831 GCACAGAAAGAGAGGCCCTCCAGCACTTCCAGTTTGTGTTCCAGGAAGAAACTTAAAGAA 3890
Qy 3642 GAAGAAACCCAGGAAGAGATGTTCTTGAACCTACACAGCAAGAGCTTACTGTGTCAGCCGTG 3701
Db 3891 CAATCAAGATGGNAGACACTTAGAGCATACAGATAAGAGGTGTGAGTGGAAACTGTA 3950
Qy 3702 CCGGTTCTGCAAAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3951 TCCATTCTGTCAAAAGACTGAGGGGACTCAAGAGCTGACAGTATGCTGTGAGAAACCC 4010
Qy 3753 GATGAGAAAAGTCAAGAGAACAGAGGTGTTTGTACACTCTGG-----ACCCAAC 3806
Db 4011 AAAGACGTACCATTTTTCGAAGGACTTGAGGGGTCTATAGACACAGGCATTAACAGTCAGT 4070
Qy 3807 AGTCAAAAGGCTGTGATGTGACATATGACAGTGAAGTGTGGAGTGGCGGGTGTGAC 3866
Db 4071 CGGGAAGAGTCACTGAAGTTGCCCTTAAGGTGAAGGACAGAAAGCTGAATGTAAA 4130
Qy 3867 GAAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGGAGGA 3911
Db 4131 AAGGATGATGCTCTTGAACCTGCAGAGTCAGCTAAGTCTCTCTCATCCCGGTGGAGAGA 4190
Qy 3912 GAGATGAAAAGTCACTTGAAGGAGAAAAGGAGAGCAAGACCCAGAGCAAGTGAAGTGAA 3971
Db 4191 GAGATGGTAGTTCAAGTTCGAAAGGGAGAAAACAGAAAGCAGAGCCCAACCCATGTGAATGAA 4250
Qy 3972 GAAGG---TGAGCAGGAAAACAGCGCTCCTGAGCATGAAGAACTTACGGGAAGCCAGTC 4028
Db 4251 GAGNAGCTTGAGCAGCAAAACAGCTGTTACCGTATCTGGAAGGTTCAGTAAGCAGCTCCTC 4310
Qy 4029 CTGACACTTCAATGCCCAGCTCAGAGAGGGGGAAGCACTGGGAAGCCTTTGGAGGAAGC 4088
Db 4311 CAGACAGTGAATGTGCCCATCATAGATGGGCAAGGAAGTCAGCAGTTTGGGAAGGAAGC 4370
Qy 4089 CTTTCTC---TCCAGACCAAGACAAAGCAGGTGTCATAGAGTTCAAGTTTCAAGCCTG 4145
Db 4371 CCTCTCCTCCCTAGCTCAAGAGGAGGCGAGTATGCAACCAAAATTTCAAGTTTCAGAGCTCT 4430
Qy 4146 GACACACAGTCACTCAAAACAGCAGAGAGCTGTGAAA-----GGTCATAGAAACCGTT 4199
Db 4431 GAGGCATCATTTCACTTAAACAGCGGCTGCAGAGGAGAAAGGTCTTAGGAGAACTGCC 4490
Qy 4200 GTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGCACTATTATACCAGCTGAG 4259
Db 4491 AACATTTTGAAGAACAGGTGAAACCGTTGAGCCCTGCAGGTGCACATTTAGTTCTGGAAGAG 4550
Qy 4260 AAGTCTCTGCAACGGGTGGCCACTGGACTTCTTCAGCATGTCAGAGGACAGCGTACCCCTG 4319

Db 4551 AANTCCTCTGAAAAAATGAAGACTTTTCCGCTCATCCAGGGGAAGATGCTGTGCCACACA 4610
Qy 4320 GGGCCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAATCTCTGCTCTGAAAGCACC 4379
Db 4611 GGGCCCGACTGTTCAGGCAAAATCGACACAGTGTAGTATCTGCTACTACCAAGAAAGGC 4670
Qy 4380 CTATCTCTGACCTACAGGAGAAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAG 4439
Db 4671 TTAAGTTTCCGACCTGGAAAGGAGAGAAAACCATCACTGAAGTGGAAAGTCAGATGAAGTC 4730
Qy 4440 GACAAGCCAGATCCTGGTCTCTGATGCTGACGGAAGGAGAGTACAGCAATCGAAAAAGTC 4499
Db 4731 GATGAGCAGGTTCCTTGCC-----AGGAGTCAAGTGTAGTGTAGCAATTG---AGAT 4781
Qy 4500 CTCGAAGCTGAACCTGAGATCTCTGGAACCTTGAGAGTAAGAGCAACAAAGATTGTGCTGAAC 4559
Db 4782 TTAGAGCCTGAAAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAACTTTGTCCAAAAC 4841
Qy 4560 GTCAATTCAGACAGCCGTTGACCAAGTTTCGCACGTAC---AGAAAAGCCCGGAAATCAT 4616
Db 4842 ATCATCCAGACAGCCGTTGACCAAGTTTGTACGTACAGAAAGAAACAGCCACCGAAATGTTG 4901
Qy 4617 GCTTATGATTTACAGACCCAGGTTT-----CTGCATGCGAGGCTTGACAG 4660
Db 4902 ACCTCTGAGTTACAGACACAAGCTCAGTGTAAAAAGCTGACAGCCAGGACGCTGGACAG 4961
Qy 4661 CAGGGAGCCCAACAGATGCTGGGCAAAAAATGAAAGATGCCAAGATGAAACACCCAGTGCC 4720
Db 4962 GAAACGGAGAAAGAGGAGGAGGAACTCAGGCCCTCTGCACAGGATGAAACACCAATTAAT 5021
Qy 4721 GCAGCCCCA-----GAGAGGACTTTCGCAAGTCTCTGACCGTTCCTGGA 4759
Db 5022 TCAGCCAAAGAGGAGTCAGAGTCAAACCGCAGTGGGACAAAGCATTTCTGTATATTTCCAAA 5081
Qy 4760 GGCATGGGCTCAGCCTCGGAAATGCTTGGCGCTTTCAGTTCGAAAGCCGCGTGTCAA- 4818
Db 5082 GACATGAGTGAAGCCTCAGAAAAAGACCATGACTGTTTGAAGTAGAAGGTTCCACTGTAAAT 5141
Qy 4819 ----- 4818
Db 5142 GATCAGCAGCTGGAAGAGTCTGCTCCCATCTGAGGAAGAGGAGGTGGAGCTGGAAACA 5201
Qy 4819 -----AGTAAGCATTTGAGAAGCTGCCT 4840
Db 5202 AAGTCTGTGCCAGAAAGATGATGGTCATGCTTGTAGCAGAAAGAAATAGAGAAGTCACTA 5261
Qy 4841 CCTCAACCCCAAGCA--TCCAAAGGAGCAGTCTGCTGATGGCCCTCAGCTCCCAAGCTTA 4898
Db 5262 GTTGAACCCGAAAGAGATGAAAAAGGTGATGATGTTGATGACCTTGAAAAACAGAACTCA 5321
Qy 4899 GCCCAGGCAGAGCCAGTGCCTCTGGAAACCTTAACCAAGAAATCCCAGACACCAACGGGA 4958
Db 5322 GCCCTGGCTGATACTGATGCTCAGAGGCTTAAACCAAGAGTCCCAGATACAAATGGA 5381
Qy 4959 CCAAGCTTAACCGAGGAGGCGCATCCCAAAAGTTGAGGTCCAGGAAGAAAGAAATGTCT 5018
Db 5382 CCAAAACAAAAAGAGAGGAGGATGCCAGGAAGTAGAATTGCGAGGAAGGAAAGTGCAC 5441
Qy 5019 ACCAAGTCAGTCAAAAGAGAAACAAGGCCCGGAGAGAGAGACCTGACGAGGACCAAGGGA 5078
Db 5442 AGTGAATTCAGATAAAGCGATCACACCCCAAGCAAGGAGGAGTTACAGAAACCAAGAGAGA 5501
Qy 5079 GACCTGGCAGAACTCCTAAGATGTTAGTTG-----CTCATTTGTAC 5117
Db 5502 GAATCTGAAAGTCAGAACTTACAGAACTTAAAAACATCATGCAAGTTAAACTCATTTGCT 5561
Qy 5118 ATCTGTAAGACCAAGAATGTGAAAAACAAAGTCACAGAAACA-----GATGCTGCTGTGGGAC 5173
Db 5562 GTTTGGAGACCAAGATGTGAACACAAAGTAGTAGAGAAATAAGTCTGCTCTGCTGAGA- 5620
Qy 5174 CTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGAGCAGGCGGCTCCAAATGATTTCC 5233
Db 5621 CTGAAGACCAAGTATTTTCAGA-ACCTTTGAGAAATTGGAGAGCAGGCAATCAACTGATCTCA 5679

QY 357 AACTCCACAGCTGTTGAAGATATCACAAAGATGGCGAGGAGGACATCAGAAATAATT 416
Db 258 AAGTCAGCGGTGTTACGACATCACAGATGATGGCAGGAGGAGACACCCGAAATAATC 317
QY 417 GAACAGATCCCTGCTTCAGAAAAAACAATGTGGAAGAAATGGTACAGCCTGCTGAGTCCAG 476
Db 318 GAACAGATTCCTTCTTCAGAAAGCAATTTAGAGAGCTAACACACCCACTGAGTCCAG 377
QY 477 GCTAATGATGTTGGCTTCAAGAAAGTATTTAAATTTGTTGGTTTAAATTCACGGTGAAG 536
Db 378 GCTAATGATATGGATTTAAGAGGTTGTTAAAGTTTGTGGCTTTAAATTCACGTGTAAG 437
QY 537 AAGGATAAAATGAAGAGTCAGATCTCAACTACTCTCACTGTCAGAGAGGTGAAGGC 596
Db 438 AAGGTAAGACAGAGAAGCTTGACACTGTCCAGCTACTCACTGTGAAGAAAGATGAAGGG 497
QY 597 GAAGGGCAGAAAGCTCTGTGCGAGCTGGAGACCAACAGAGGCCAGTGTGGAGACTGCC 656
Db 498 GAGGAGCAGCA-----GGGGCTGGCGACCAAGAGNCCCGACCTTGGGGCT--- 545
QY 657 GTCGAGAGTCAGATCCAAAGAAAGTGAGCTGAAGCAATCCACAGAGAGCAAGAGGC 716
Db 546 ---GGAGAAGCAGCATCCAAAGAAAGCGAAACCCAAAACAATCTACAGAGAAACCCGAAAG 602
QY 717 ACCCTGAAGCAAGACAGACGACACAGAAATCCCCCTTCAAGCCGATCTGATCAAGCG 776
Db 603 ACCCTGAAGCGTGAGCAAGGCCAGCAAGAAATTTCTCCCCCGCCGAATCTGGCCAA--- 659
QY 777 GCTGAGGAAGAACCAAGATGAAGGAGAGAAAGAAACAGAGAAAGAGCCACCAAGTCC 836
Db 660 GCAGTGGAGGATGCAAGAGGAGGAGAGAGAGAAACAGAAAGAAACCTAGCAAGTCT 719
QY 837 CCAGAAATCCCGAGCAGCCAGTCAACAGTGAGCAACATCTTCTTCAAGAAAGTCTTC 896
Db 720 GCAGAAATCTCCGACTAGTCCGCTGACCAGTGAACAGGATCAACCTTCAAAAAATCTTC 779
QY 897 ACTCAGCGTTGGCGCGCTGGCGCAAGACACAGCTTCAAGAAATCAAAAGAGGATGAT 956
Db 780 ACTCAGGTTGGCGCGCTGGCGCAAAAGACCAAGTTTCAGGAAGCCGAGAGGATGAA 839
QY 957 CTGGAATCTCGAGAGAGAAAGAGCAAGAGCGCAAAAGATPAGACGAGGAGAAAG 1016
Db 840 GTGGAGCTTCAGAGAAAGAAAGGAACAGAGCCAGAAAGATGACACAGAAAGAGAC 899
QY 1017 GAAAGACAGAGCCAGCTCCGAGGA-----GAGGAG 1049
Db 900 GGAAGGCAGAGGTTGCTCCGAGAAACTGACCGCTCCGAGCAAGCCACCCACAGAG 959
QY 1050 CCGCAGAGACACAGACCAAGGCCAGGTTGTGACGACACTACAGAAAGGTGGAGTGCCT 1109
Db 960 CCGCAGAAAGTGCCACAGAGCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCC 1019
QY 1110 TTGGAAGACAGGTTGTGACCTGGAGGACATCGTCAGAGAGAGTGTGCTCTTTGGCA 1169
Db 1020 TCAGAGAGCAAGTCACTAGTGGCTCGAGGACCTTCTGAAGAGAAACCTGCTCGGTTGGCG 1079
QY 1170 ACGGAAGTGTGATGAGAAAGTGAAGCCCAAGAA---GTTGTTGAGAGGTCAC 1226
Db 1080 ACAGAAAGTGTGATGAGAAATAGAGTCCACCAAGAGAGGTTGTGSCCGAAGTCCAC 1139
QY 1227 GTGAGCAACCGTGAGAAAGACAGAGGAGGAGGAGGAGGAGAGGAGGAGGAGGAGG 1286
Db 1140 GTCAGCACCGGTGGAGGAGAGAACCGAAGAGCAGAAA----- 1175
QY 1287 GTGTGCTAGAGAAACAGAGAGATCTTCCCTCTGAGAACTGGCTGAGCCGAGAG 1346
Db 1176 ACGAGGTGAAGAAACAGCAGGGTCTGTGCCAGCTGAAGAAATGGTTGAATGGATGCA 1235
QY 1347 GTCCCCCAGGAAGCTGAGCTGTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCT 1406
Db 1236 GAACCTCAGGAGCTGAACCTGCCAGGAGCTGTGAGAGCTCAAGAAACGTGTGTTCC 1295
QY 1407 GGAGGAGACCACACTCAACTGACAGACCTAAGTCTCTGAAGAGAGAGCGCTGCCAAAAC 1466

Db 1296 GGAGAGACCCCTTACACAGGAGCTGACCTCAGTCCTGATGAGAAGGTGCTGTCAAACCC 1355
QY 1467 CCAGAAAGGCATTTCTAGTGAAGTGGAGATGCTCTCTCTCAGAAAAAGAAATCAAGGTACAG 1526
Db 1356 CCCGAAGGCGTTGTGAGTGAAGTGGAAATGCTGTCTCATCAGAGAGAGATGAAGTCCAG 1415
QY 1527 GGAAGTCCCTTGAAGAAACTTTCTAGTGAAGTCTAGGCTTAAAGAGCTGTCTGGGAAGAG 1586
Db 1416 GGAAGTCCCACTAAAGAAAGCTTTTACAGCACTGGCTTAAAAAAGCTTTCTGGAAGAAA 1475
QY 1587 CAGAAAGGGAACAGAGAGTGGGGAGAGAGAGCCTGGAGAAATACCAACATTCAC 1646
Db 1476 CAGAAAGGGAAGAAAGGA---GGAGGAGACAGGAATCAGGGAGACACTCAGGTTCCA 1532
QY 1647 ACCGAATCCCCACAGAGTGTCTGATGAGCAGAGGAGAGAGCTCTGCTGCTCCCCCGAG 1706
Db 1533 GCCGATTTCTCCGACAGCCAGGAGGAGCAAAAGGCGAGAGCTCTGCCTCATCCCCTGAG 1592
QY 1707 GAGCCTGAGGAGACCACTGTCTGGAGAAAGGGCCGCTGGAAGCACCCAGGATGGGAA 1766
Db 1593 GAGCCCGAGGAGATCACTGTCTTGGAAAGGGCTTAGCCGAGGTGCGAGAGATGGGAA 1652
QY 1767 GCTGAGAGAGGAATCACTTCCGATGAGAGAAAGAGAGAGGAGTCACTCCCTGGCA 1826
Db 1653 GCTGAAGAGGAGGACTACTTCCGATGAGAGAAAAAAGAGAGAGGTGTCACTCCCTGGCA 1712
QY 1827 TCCTTCAAAAAAGATGTGACCAAGAAACGGGTCCGAAAGCTTCTGAGAGTGACAAG 1886
Db 1713 TCATTTCAAAAAAGATGTGACGCCCAAGAGCGTGTAGACGGCTTCGGAAGTGATAA 1772
QY 1887 GAGAAAGAGCTGGAGAGGTCAAGAGCGCACTTGTCTCTCACTGATAGCACAGTGTCA 1946
Db 1773 GAAGATGAGCTGGACAAGGTCAAGAGCGCTACCTGTCTTCCACCGAGAGCACAGCCTCT 1832
QY 1947 GAATGCAAGATGAAGTCAAACTGTTGTGAGGAACAAAGCCAGAGAACCAAGCGT 2006
Db 1833 GAAATCAAGAAAGAAATGAAGGGAGCGTGGAAAGCCAAAGCCGGAAGAACCAAGCGC 1892
QY 2007 AGGTGGATACTTTCAGTGTCTTGGGAAGCACTGATTTGTGTCGATCATCTCAAGAGAGA 2066
Db 1893 AAGTGGATACCTCAGTATCTTGGGAAGCTTTAATTTGTGTGGATCATCCAAGAAAGA 1952
QY 2067 GCAAGAAAGCATCTCTTTCAGATGATGAAGAGGCGCAAGGACACTGGGAGGGGACAGT 2126
Db 1953 GCAAGGAGAGGTCCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGGAGAGACCAC 2012
QY 2127 CACAGAGCAGAGGAGGCGAGCAAGAAAGAACCGGAAACAGACGCTGTTCTCTGCAGC 2186
Db 2013 CAGAAAGCTGATGAGGCGCGAAAGAACAAAGAGACGCGGACAGACGGGATCTTGTGCT 2072
QY 2187 ACCCAGAGCAGGACCAAGCGCAAGGAAGTTCTCTACCCGAGCCAGCGGAAGCCCTTCC 2246
Db 2073 TCCAAAGAACATGATCAGGGCAGGGAAGTTCTCTCCCGAGCAAGCTGGAAGCCCTACC 2132
QY 2247 GAAGGGGAAGGTGCTCTCACCTTGGGAGTCAATTTAAAGATTTAGTCACTCCAGAAAAA 2306
Db 2133 GAAGGGGAGGCGCTTCCACTCGGAGTCAATTTAAAGGTTAGTCAACGCCCAAGAAAAA 2192
QY 2307 TCCAGTCAAACTGGAAGAGAAAGCCGAGAC-----TCTAGTGTAGAGCAG 2354
Db 2193 TCAAGTCCAAAGCTGAAGAGAAAAAGCAAGACTCCATAGTGGGTCTGGTGTAGAACAT 2252
QY 2355 TTGTCCACTGAGATCCGAACCGAGTAGAGAGATCTTGGGTTTCCATTAAGAAATTCATC 2414
Db 2253 TCCACTCCAGACACTGAACCCGGTAAAGAGAAATCTCGGTCTCAATCAGAAAGTTATT 2312
QY 2415 CCCGGAAGGCGGAGAAAAAGGGCAGACGGGAGCAAGCAAGCAAGCCACTGTGGAAGACTCA 2474
Db 2313 CTTGGAAGGAGGAGAAAAAGGGCAGATGGGAACAGAAACAGCCCTCTGTTGAAGCGCA 2372
QY 2475 GGCCCAAGTGGAGATAAATGAGGACGACCTTAATGTGCCAGCGCTCGTCTGTCTGAG 2534

Db 2373 GGGCCAAAGGGGCCAAGAAATGACTCTGATGTCGCCGCCGTGTCCCTCTGTCTGAG 2432
Qy 2535 TATAATGAGTGGAGAGGAGAGATGG-----AAGCCAGGGGAAATCGAGCTG 2585
Db 2433 TATGATGCTGTAGAAAGGAGAAATGGAGGACAGCAAGCCCAAAAAAGCCAGAGCAG 2492
Qy 2586 CCCCAGCTGCTGGGGGCTGTGATCGTGTCCGAGGAGCTCAGTAAAGACTGTGTCCACA 2645
Db 2493 CCGGAGCAAGAGGACAGCACTGAGGTGTCCAAAGAGCTCAGGAGATCAGGTTCAATATG 2552
Qy 2646 GTGATGTGCGAGTCATTGATGGAGCCAGGGCAGTCCACGATGTGCAAGAGCGGTCTCT 2705
Db 2553 ATGGAGAGAGCTGTGCTGCTGATGGAGGAGGAGGAGCAGCTATTATGAAAGAGTCTCCT 2612
Qy 2706 TCGTGGATATCCGCTTCCGTTACAGAACTCTTTGAACACACAGCGGGAGAACCTGCCA 2765
Db 2613 TCTTGGATATCTGCTTCAGTGCAGAACTCTTTGAAACAAGTGAAGCTGMAACCCACATG 2672
Qy 2766 CCTGTGAAGAGTCACTGMAAAAGACATCTTGC---AGAGAAACTCTGTGCTCAC 2822
Db 2673 TTAACTGAGAGGTATTTGAAAAGAGAAATTTGCAAGAAAGAACCCCCACGGTTACT 2732
Qy 2823 CAGACGTTACAGAGGGGTAAAGATGCCATGACGACATGTGACACAGTGAAGTGAATTTC 2882
Db 2733 GAACCTCTGCCAGAAACAGAGAGGCCCGGGGCGACACGGTCTGTTAGTGAAGCGGAATTG 2792
Qy 2883 ACCTGAGAACTGTGACAGACCAAGAGACCTCAGAGGCTCTCCGTACTGAAGAACTTACC 2942
Db 2793 ACCCCGAAAGCTGTGACAGCTGCAGAAACTGCAAGGCCACTTGGGTGCCGAAAGAGAAC 2852
Qy 2943 GAAGCATCGGGGGCCGAGAAAGACACAGCATGTGTCCGCAATTTCCAGCTGACTGAC 3002
Db 2853 GAAGCATCTGCTGTGAAGAGACACAGAAATGTGTGACAGTCTCCAGTTAACCGAC 2912
Qy 3003 TCCCAGACACCAAGAGAGAACCCCGAGTTCAGAGAGTGAAGAGTGTGTGATGAT 3062
Db 2913 TCCCAGACACCAAGAGAGAGGCTCCGCTGTCAGAGAGTGAAGTGTGCTGATGAC 2972
Qy 3063 ACAGAAAGAGAGAGCGCCAGACGAGGCCATCTTCAAAGCCGTTCAGACAAAGTGAAA 3122
Db 2973 ATAGAAGAGCAAGAGAGGCGGACTCAAGAGTCTTCAAGGAGTGCAGAAAAAGTGAAA 3032
Qy 3123 GAGGAGTCCCAAGTGTCTTCAAC---CAGACTGTGCAAGAAAGGGGTCAAAAGACTG 3179
Db 3033 GAGGAATCCCACTGTGCTGCAACCGGTGGCCAGAAAGTGTCTTCAAGCTGTGCAGAA 3092
Qy 3180 GAGAAGTGTGAAGAGTGAAGAGGACTCCGAAGTCTGCTTCGAGAAAGAGAGAC 3239
Db 3093 GCAGAGGCAAGAAAGACCAAGAAAGAGGCTGAAGCCTCGGTCTGAAGAAAGAGCGAT 3152
Qy 3240 GTTATGCCGAAAGGACCCGTGACGAGAGTGAAGTGAAGTGTGACATCTTGCACAGGGCTCTGAG 3299
Db 3153 GTAGTGTGAAAGTGAATGATCTCAGAGGCAAAAACTGAGCTTTTTCACAAGGGAAAGTGTG 3212
Qy 3300 ACTGGAACAGGCTACTTCAAGAGGCTTTGAAA---GTTCCGAAAGTCAAGCAAGTGAAGAC 3356
Db 3213 GTGGGGCAGACCAACCCCAAAAGCTTTGAAAAAGCTTCAAGTCAAGAGAGCAATGAG 3272
Qy 3357 -----CATGTGCCAGTGTCCA-----GGTTATCAAGCTCCAG 3389
Db 3273 TCCAGTGAAGCTTTGAACCACTTGTCAAGCCGAACCTTAGCTGGGGTAAAAATCACAGGAG 3332
Qy 3390 CAGCTGATGAAGACAGGCGCTGGCCCTGAGTCAATCCGAAACCTTGAACAGACAGTGAAGAC 3449
Db 3333 ATGGTATGTAAGAACAGGCTATCCCCCTGACTGGGTGAAAACCCCTACAGACAGTGAAGCT 3392
Qy 3450 AATGGAAGCACTCTCTTACAGATTCAGACACTGCAAGTGGGACACAGCAAGTGAACCC 3509
Db 3393 GATGGAAGCAACCCCGTACCCGACTTTGACGCAACAGGACCAACCCAGAAAGCGAGATT 3452
Qy 3510 AATTGACAGCAGGACAGTAAAGCACTGCAAGCTGTCAAGGAGTCAAGGTCACAGAGAA 3569
Db 3453 GTGGAATCATGAGGAATGAGGTGCAATCTGTATCCAGTCAAGGGGCAAGAAAGCA 3512

Qy 3570 GAGGCGCTACTGCTCAAGAAAGAGAGCTTTGACACTACTAATTAATGTTCCAGCCAG 3629
Db 3513 GAGGCGACTCTTGCACAGAAAGAGAGGCTTCAGACCTTCAAGTGTGTTGTTCCAGAAA 3572
Qy 3630 GAAGAACATGGGGAGAAACCAAGAAAGATGTTCTTGAACCTTACACAGACAGACTTACT 3689
Db 3573 GAACATAAGAACAAATCAAGATATGGAAGACACTCTAAGAGCATATACAGATTAAGAGTGTCA 3632
Qy 3690 GCTGACCGTGGCCCGTTCTGGCAAGAATGAGTGGGTCAAGAGGTTAGTGAATGG 3749
Db 3633 GTGGAACCTGTATCAATCTGTCAAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCT 3692
Qy 3750 TTGATGGAAGAAAAGTCAAA-----GAAGAACAGAGGTTTGTCACTGTGG- 3799
Db 3693 GATGGAAGAACCAAGAGATGACATTTTTCAGAGGACTTGAAGGTTCTATGACACAGGC 3752
Qy 3800 -----ACCCAAAGTCAAAAGCTGTGATGTGACATATGACAGTGAAGTGAAGTGGAGTGG 3854
Db 3753 ATAAAGTCACTCGGAAAGAGTCACTGAATTTGCCCTTAAGGTGAAGGACAGAAAGAA 3812
Qy 3855 GCCGGGTGTCAAGAAAAGAGAGTACTGAAGTCAAGCTTTAG-----C 3899
Db 3813 GCTGAATGTAAAAAGATGATGCTCTTGAACCTGACAGAGTCAAGTAAAGTCTCTCATCC 3872
Qy 3900 CTGAGAGGAGAGATGAGAACTGACGTTGAAAAGAGAAAGAGGACAAAGCCAGAG 3959
Db 3873 CCGGTGAAGAGAGATGATGTTCATGTCAGAAAGGAGAAACAGAAAGCAGAGCAACC 3932
Qy 3960 CAAGTGAAGTGAAG---AAGGTGAGCAGGAAACAGCCGCTCTCTGACATGAAGAACTTAC 4016
Db 3933 CATGGAATGAAGAGAAAGCTTGAAGCAGAAACAGCTGTATACGTATCTGAAGAGTCAAGT 3992
Qy 4017 GGAAGCCAGTCTTGAACATTTGAACATGACCTCAGAGAGGGGAGAGCACTGGGAAGC 4076
Db 3993 AAGCAGCTCTTCCAGACAGTGAATGTGCCATCATATGATGAGCAAGAGAGTCAAGCAGT 4052
Qy 4077 CTGGAAGAGACCTTCTTC---TCCAGACCAAGACAAAGAGGTGTGATGAGTTCAT 4133
Db 4053 TTGGAAGAGAACCTTCTCTCTGCTGCTAGGTCAAGAGGAGGCAAGTATGCAAAATTTAA 4112
Qy 4134 GTTCAAGCTTGAACACACAGTCACTCAACAGACAGAGTGTGAGAAA-----GGTTC 4187
Db 4113 GTTCAAGACTCTGAGGCAATCATCTTAACAGCGGCTGACAGAGAGAGAAAGTCTTA 4172
Qy 4188 ATGAAGACGTTGTGATTTCAAGAGCAGGTGAAGTCCAGAGTGTGATGTCACATTTA 4247
Db 4173 GGAGAAACTGCAACATTTTGAAGAAACAGGTGAACGTTGAGAGCTGCAAGTGCACATTTA 4232
Qy 4248 TTACCAAGTGAAGATCTCTGCAACCGGTGGCCACTGGAATCTTCAACATGCAAGAGAC 4307
Db 4233 GTTCTGAAGAGAAATCTCTGAAAAAATGAAGCTTTGCGCTCATACAGGGGAAAGAT 4292
Qy 4308 ACGGTACCCCTGGGGGCTGAGTCTCAGCAGAAATCATCCCAATCATATGTAATCTCTCT 4367
Db 4293 GGTGTGCCACAGGGGCCGACTGTCAAGCAAAATGACACAGATGATGTAATCTCTACT 4352
Qy 4368 CCTGAAGACACCTTACATCTGACCTTCAAGAGAAATPAAGCGATCCCAAGAGAGCGA 4427
Db 4353 ACCAAGAAAGGCTTAAAGTTTCCGACTTGAAGAGAGAAACCAACATCTGAAGTGAAG 4412
Qy 4428 TCAAGAGAAAGAGACAAAGCCAGATCTGTCTGATGCTGACAGGCAAGAGATGACAGCA 4487
Db 4413 TCAGATGAAGTGAAGACAGAGTGTCTTGGC-----AGAGAGTCAAAAGTGAAGTATAC 4466
Qy 4488 ATCGAAAAAGTCTCAAGGCTGAACCTGAGATCTGGAATCTGAGATGAAGTGAAGCAAG 4547
Db 4467 ATTTG---AGATTTGAGCTGAAGATGGAATGGAATTTGAACTTGAACCAAAAGCGTAAA 4523
Qy 4548 AATTGTGTAAGTCAATTCAGACAGCCGTTGACCAAGTTGCAAGCTTAC---AGAAAAGCC 4604
Db 4524 CTTGTCAAAAATCATCATCAAGACAGCCGTTGACAGTTTGTATCGTACAGAAAGAAACAGCC 4583

Qy	4605	CCGAAACTCATGTTTANGATTACAGACCAGGGTTC-----TGCAATG	4648
Db	4584	ACCGAAATGTGTGCGCTGAGTTTACAGACAAGCTCACGTGATAAAGCTGCACGCCAG	4643
Qy	4649	CAGCCTTGACAGCAGGAGCCCCAACAGATGCTTGGCAAATAATGAAGAATGCCAAGTAGAA	4708
Db	4644	GACGCTGGACAGGAACCGGAGAAGAAGAGAGAGAACCTCTGGCCTCTTGCAACAGATGAA	4703
Qy	4709	ACACCCAGTCCCGCAGCCCC-----GAGAGAGCTTGTCAAGTCCT	4747
Db	4704	ACACCAATTACTTTCAGCCAAGAGGAGTCAGAGTCAACCGCAGTGGACAAGACATTTCT	4763
Qy	4748	GACCGTTCTGGAGCATGGGCTCAGCCTCGBAATGCTTGGCCGCTTGCAGATTGTAAGC	4807
Db	4764	GATATTTCCAAGACATGAGTGAAGCCTCAGAAAAAGACCATGACTGTTTGAAGTAGAAGT	4823
Qy	4808	GC CGGTGTC AA-----4818	
Db	4824	TCCACTGTAATGATCAGCAGCTGGAAGAGTGTCTCCCATCTGAGAAAGAGGAGGT	4883
Qy	4819	-----AGTAAGCATTT	4828
Db	4884	GGAGCTGGAACAAGTCTGTGCCAGAAGATGATGTCATGCCCTTGTAGCAAGAAAGATA	4943
Qy	4829	GAGAGCTGCTCTCAACCCAAAGA--TCCAAAAGGAGCAATGCTGTGATGGCCCTCAG	4886
Db	4944	GAGAAGTCACTAGTTGAACCGAAGAAGATGAAAAGGTGATGATGTTGATGACCTGAA	5003
Qy	4887	CTCCAAGCTTTAGCCCAGCAGGAGCCAGTGCCTCTGGNAACCTTAACCAAGAATTCGCCA	4946
Db	5004	AAC CAGAACTCAGCCCTGGCTGATCTGATGCTCAGGAGGCTTAAACCAAGAGTCCCCA	5063
Qy	4947	GACACCAACGGACCAAAAGCTTAAACCGAGGAGGCGATCCCCCAAAAGTTGAGTCCAGAA	5006
Db	5064	GATACAAATGACACNAAA CAANAAGAGAGAGAGGATGCCAGGAAGTAGAATTCGAGAA	5123
Qy	5007	GAGA AATGTCTACCAAGTCAGTCA AAGAGAA CAAGGCCCGCAGGCAAGAGAGACCTCAG	5066
Db	5124	GGAAAAGTGACAGTGAATCAGATAAAGCGATCACACCCCAAGCACAGGAGGAGTTACAG	5183
Qy	5067	GAGCCAAAGGAGACCTGGCGAGATCCTTAAGATGTTAGTT-----5106	
Db	5184	AAACAAGAGAGAGAATCTGCMAAGTCAGAACTTACAGAACTCTTAAACAATCATGCAGTTA	5243
Qy	5107	-GCTCATTGTACATCTGTAAGACCAAAATGTGAAAACAAGTCACAGAACAA---GATGC	5161
Db	5244	AATCATTGTCTGTTTGGAAAGACCAAGATGTGAGACAAGTAGTAGAAGAAATGAATGC	5303
Qy	5162	TGCTGTTGGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGACGAGGCCGT	5221
Db	5304	TGCTGTCTGAGA-CTGAAGACCAAGTATTTTCAGA-ACTTTTGAGAATTTGGAGAGCAGGCACAT	5361
Qy	5222	CCAATGATTTCCACCCAGTAGACGACCCCGACAATTCGTGAGGCTTCATCGGGAGCTAGAG	5281
Db	5362	CAACTGATCTCATTTCTTAGAGAGC-CCCTGACAATCCTGAGGCTTCATCAGGAGCTAGAG	5420
Qy	5282	CCAGCTAAACATTTCTCGTTTTCAAGACTGCTTTTGATTTGCCCTTGTATGCGGTCCTGCT	5341
Db	5421	CCATTTAAACATTTCTCTTTCCAAAGACCAACCTTACAATTTCCCTTGTATAAC-----5472	
Qy	5342	ATTTCTAAACATTTCTCGTTTTCAAGACTGCCTTTTGATTTTGCCCTTGTATGCGGTCCTGCT	5401
Db	5473	-----CATATA	5478
Qy	5402	ATTTGGAATTTAAGTCTCGGTTCTCAACCTGGAACCA-ATTCTGCCAATACCTAGTTCC	5460
Db	5479	AATTTCAITTTAAGGTCCTAAATTTCTTAACCTGGAACTGGAGTTGGCAATACCTAGTTCT	5538
Qy	5461	ACTTCTCAAACTGGAGCATCCTCCTTTATGTAATTAATATGATGTTTATGTAAGTCTCTCC	5520
Db	5539	GC TTCTGAAACTGGAGTACATTC TTTACATATTTATGATGTTTAAAGTAG---TCC	5595
Qy	5521	TCCTGTACCTATTGTATATTTTTTTTCTTAACGTTT-----AAGCACATG	5563

Db	5596	TCCGTATCTATTGTTATATTTTTTCTTTAAATGCTTTAAGGAATGTGCAGGATACTACATG	5655
Qy	5564	CTTTTGTGTAATTGCAATATATAACGGGTGTGCAGCCATAGCAGCGCTTCGAAAGCTCC	5623
Db	5656	CTTTTGTGATCACACAGTATATGATGGGGCATGTGCCATGTGCAGCGCTGGGAGCTTT	5715
Qy	5624	AAGCCTCAACTGTAACTGTGCAGCAACAAGATAAACATTC-----CTGGCAAGAAGAG	5674
Db	5716	AAGCCTCAGTTATATAAACCCACGAAAAACAGAGCCTCTAGATGTAAACATTCCTGATCAA	5775
Qy	5675	ACAAGTCTTTTTTAAAAGTTTACTGATGCTTAGATCTGTGGGCTCTAGTCTCTCGAAAGT	5734
Db	5776	GGTACAATTCCTTTAAAATTCATAATGATCAGAGTGCCAATTTTAACTGTGCTACTCTGAAAT	5835
Qy	5735	GGTGTGTTTCTATGACAGCGAGCTCAGAAATAAAAAACCCCATTTTCAAAACATCCAGGA	5794
Db	5836	GGTCACTTTCCTATTACAGGAGTGTCTAAAACATAAAAGCATTTTGAACATACAGAA	5895
Qy	5795	TGTCCCAATATACCATGATTTTTTCCCCCTTTTGTCTAATCCAGTCCAGTTGGAAAG	5854
Db	5896	TGTTCTAATGTCAATTGGGAAATTT-----TTCTTTCTAACCCAGTGGAGGTTAGAAAG	5948
Qy	5855	AGTCTCTCTGTGTCCAGA-----TTAAGCCCTGTCTCTTTAATGATAT	5897
Db	5949	AAGTTATATTCTGTGTAGCAAAATTAACTTACATCCTTTTCTACTTGTATTGTTGTTT	6008
Qy	5898	GGACAAATGAGTGTGCCCTTAAGGCCATGAGA-----TGTTTCTCTAATGCA	5941
Db	6009	GGACCGATAAGTGTGCTTAATCCTGAGCAAGTAGTGAATATGTTTATATGTTATGAA	6069
Qy	5942	GAAAGAACTGTGTGACGTTTTTTTTTGATGTACTCTCTATGCTGGACCGAATTCATATG	6001
Db	6069	GAAGAAGAAATGTTGTAAG- TTTTGTGATTCTACTCTTATATGCTGGACTGCATTACACA	6126
Qy	6002	CAGATCGAAGTGTGCTGTTCTTTTACAGATGGTATTTTGTATAGATACCTGAGTTGTCT	6061
Db	6127	TGGCATGAAATAAGTCAGGTTCTTTTACAAATGGTATTTTGTATAGATACCTGATTGTGTT	6186
Qy	6062	GTGTTATATCTGTGCCCCCTCTTTTAAAGAACAAATGTTGCATTATGTTCTCTTTGGATAAAAT	6121
Db	6187	GTGCCATATTGTGCCATTCTTTTAAAGAACAAATGTTGCAACACATTCATTTGGATAAGTT	6246
Qy	6122	GTGATTTGACAACATGATTTAAATAAA	6147
Db	6247	GTGATTTGACCACTGATTTAAATAAA	6272

[illegible]

Db 10693 CAGAAAGGAAAGAGGA---GAGGAGACGAGGAATCAGGGAGAGCACTCAGGTTTCCA 10749
 QY 1647 ACCGAATCCCAGAGAGTGTGTAGTACAGAAAGGAGAGAGTCTGTGCTGTCCCCCGAG 1706
 Db 10750 GCCGATTTCTCCGACACGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCTCTCATCCCCCTGAG 10809
 QY 1707 GAGCCTGAGGAGACCAAGTCTCTGAGAGAAAGGCCGCTGGAAGCACCCAGGATGGGGAA 1766
 Db 10810 GAGCCCGAGGAGATCAGCTGTCTGGAAGAGGGCTTAGCCAGGTGACAGAGGATGGGGAA 10869
 QY 1767 GCTGAGGAGAACTACTTCCGATGGAGAGAAAGAGAGAAAGGATCACTCCCTGGGCA 1826
 Db 10870 GCTGAAGAGGAGTACTTCCGATGGAGAGAAAGAGAGAGGTGTCACTCCCTGGGCA 10929
 QY 1827 TCCTTCAAAAGATGTGTACACCCCAAGAAACGGGTCCGAAGACCTTCTGTAGAGTGTACAAG 1886
 Db 10930 TCATTCAAAAGATGTGTAGCCCAAGAACGCTGTAGACGGCTTCGGAAAGTGATATA 10989
 QY 1887 GAGGAAGCTGAGAGAGGTCAAGAGCGCCACTTGTCTCTCACTGTATAGCACAGTGTCA 1946
 Db 10990 GAAGATGAGTGTGACAAAGTCAAGAGCGCTACCTTGTCTTCCACCGAGAGCACAGCCTCT 11049
 QY 1947 GAAATGCAAGATGAAGTCAAAACTGTTGTGAGAAACAAAGACAGAGAACCAAGCGT 2006
 Db 11050 GAATGCAAGAGAGATGAAGGGAGCGTGTGAAGAGCAAAAGCCGGAAGAACCAAGCGC 11109
 QY 2007 AGGTTGGATATTTCACTGTCTTGGGAGCACTGATTTGTGTGCGATCATCCAAAGAGAGA 2066
 Db 11110 AAGTTGGATACCTCAGTATCTTGGGAGCTTTAATTTGTGTGGATCATCCAAAGAAAGA 11169
 QY 2067 GCAAGGAAGCATCTCTTTCAGATGATGAAGAGGGCCAAAGACACTGGGAGGGGACAGT 2126
 Db 11170 GCAAGGAGAGGGTCTCTTCTGTATGAGGAAGGGGACCAAAAGCAATGGGAGGAGACCAC 11229
 QY 2127 CACAGAGCAGAGGAGCCACCAAGACAAAGAACGCGGACACAGCCTGTTCTGTCCAGC 2186
 Db 11230 CAGAAAGCTGATGAGGCCGGAAGAAACAAAGACGGGGACAGCGGATCTTGTGTGT 11289
 QY 2187 ACCGAGAGCAGACCAAGCCGCAAGGAAGTTCTCTCACCCGAGCCAGCGGAAGCCCTTCC 2246
 Db 11290 TCCCAAGAACATGATCAGGGCAGGGAAGTTCTCTCCCGGAGCAAGTGGAGCCCTACC 11349
 QY 2247 GAGGGGAAGGTGTCTCCACTTGGGAGTCAATTTAAAGATTAGTCACTCCAAAGAAAAA 2306
 Db 11350 GAAGGGGAGGGCGTTTCCACTTGGGAGTCAATTTAAAGGTTAGTCAGCCCAAGAAAAA 11409
 QY 2307 TCCAAGTCAAACTGGAAGAGAACCGGAGAC-----TCTAGTGTAGAGCAG 2354
 Db 11410 TCNAAGTCCNAGCTGGAAGAGAAAGCGAAGACTCCATAGCTGGGTCTGGTGTAGAACAT 11469
 QY 2355 TTGTCCACTGAGATCGAACCCGATAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATC 2414
 Db 11470 TCCACTCCAGACACTGAACCCGTAAGAAGATCTGGGTCTCATCAAGAGTTTATT 11529
 QY 2415 CCGGAGCGGGGAAGAAAGGGCAGACGGGACAGAAACAGCCACTGTGGAGACTCA 2474
 Db 11530 CTTGGAAGGAGGAGAAAGAGCCAGATGGGAAACAGAACAGCCCTGTTGAAGACGCA 11589
 QY 2475 GGGCCAGTGGAGATAATGAGGAGCACCCTAATGTCTCCAGCGTCTGTCTGTCTGAG 2534
 Db 11590 GGGCCAAACAGGGGCCAACGAAGATGACTGTGATGTCTCCGCGCGTGTCTGTCTGAG 11649
 QY 2535 TATAATGCAAGTGGAGGGAGAGATGG-----AAGCCCGAGGGGAATACGGAGCTG 2585
 Db 11650 TATGATGCTGTAGAAAGGAGAGAAATGGAGGACACAGCCCAAAAGCCGAGAGCAG 11709
 QY 2586 CCCCAGCTGTGGGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGTGTCACACT 2645
 Db 11710 CCGAGCAGAGGACGCCACTGAGGTGTCCAAGGAGCTCAGCGAGAGTCAGGTTTCATATG 11769
 QY 2646 GTGAGTGTCCAGTCAATTGATGGGACCGGGCAGTCAACAGTGTCAAGAGCGGTCTCT 2705

Db 11770 ATGGCAGCAGCTGTCTCCTGACGGGACGAGGGCAGCTTACCATTTTGAAGAAAGGTCTCCT 11829
 QY 2706 TCGTGATATCCGCTTTCGGTAACAGAAACCTCTTTGAACACACACGCGGAGAGCATGCCA 2765
 Db 11830 TCTTGATATCTGCTTCAGTGTACAGAACTCTTTGAACAGTAGAAGCTGAAGCCGACTG 11889
 QY 2766 CTTGTTGAAGAGGTCTACTGAAAAGAGATCATTTGC---AGAAGAACTCTCTGTGCTCAC 2822
 Db 11890 TTAACCTGAGGAGGTATTGGAAGAGAAAGTAATTGCAAGAAAGAAACCCCCACGTTACT 11949
 QY 2823 CAGACCTTACCAGAGGTAAAGATGCCATGACATGTCACCATGTCACCAAGTGAAGTTC 2882
 Db 11950 GAACCTCTCCAGAGAACAGAGAGGCCCGGGGGGACACCGTCTGTAGTGAAGCGGAATTG 12009
 QY 2883 ACCTCAGAAGCTGTGTACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAAGAAATTAC 2942
 Db 12010 ACCCCGAAAGCTGTGACACTGCAGAACTGCAGGGCCATTGGGTGCCGAAGAGAAC 12069
 QY 2943 GAAGCATCGGGGCCGGAAGAGACCAAGACATGTGTCTCGCAGTTCCTCCAGCTGACTGAC 3002
 Db 12070 GAAGCATCTGCTGTGAAGAGACCAAGAAATGTGTCTCAGCAGTCTCCCAAGTTAACCGAC 12129
 QY 3003 TCCCCAGACACACAGAGGAGAACCCACAGTTTCAGGAGGTAGAGTGTGTCTAGAT 3062
 Db 12130 TCCCCAGACACACAGAGAGGAGCCACTCCGGTGCAGAGGTGGAAGGTGGCTACTGAC 12189
 QY 3063 ACAGAAGAGAGGAGCGCCAGACGAGGCGCATCTCCAAAGCCGTTGCAGACAAAGGTGAA 3122
 Db 12190 ATAGAAGACAGAGAGGCGGACTCAAGAGTCTCTCCAGGAGTGCAGAAAGTGA 12249
 QY 3123 GAGGAGTCCAGGTGCTGCAACC---CAGACTGTGCAGAGAACCGGGTCAAAAGCACTG 3179
 Db 12250 GAGGAATCCAGCTGCTGGCACCGGTGGCCAGAAAGATGTGCTTCAGCCTGTGCAGAGA 12309
 QY 3180 GAGAGTTGAGGAGGTAGAGGAGTCCGAAGTGTCTGGCTTCGAGAGAAAGAGAGGAC 3239
 Db 12310 GCAGAGGCAGAAAGACCAAGAGAGGAGGTGAAGCGTCTGGGTCTGAAGAAAGAGACGGAT 12369
 QY 3240 GTTATGTCCGAAAGGAGCCCGTGCAGGAAGCTGGAGCTGAGCATCTTCACAGGGCTCTGAG 3299
 Db 12370 GTAGTGTGAAAGTAGATGCTCAGGAGGCAAAACCTGAGCCTTTACACAGGGAAGGTG 12429
 QY 3300 ACTGAGACGGCTACTTCCAGAGAGCCCTTG---AAGTTCTGAAGTCAAGGACAGATGTA--- 3353
 Db 12430 GTGGGGCAGACACCCAGAAAGCTTTGAAAGAGCTCTCAAGTCAACAGAGAGCATAGAG 12489
 QY 3354 -----GACCATGTCCGACAGTGCCA-----GGTTATCAAGCTCCAG 3389
 Db 12490 TCCAGTGAAGCTTTGTAACCACTTGTCAAGCCGAAACCTTAGCTGGGGTAAATTCACAGGAG 12549
 QY 3390 CAGCTGATGGAAACAGGCGCTGGCCCTGAGTCACTCCGAAACCTTCAGACACAGTGAAGCA 3449
 Db 12550 ATGGTGTAGGAACAGGCTATCCCCCTGACTCGGTGGAAACCCCTTACAGACAGTGAATCT 12609
 QY 3450 AATGAAGCACTCTCTTAGCAGATTCAGACACTGCAGATGGGACACAGCAAGATGAAACC 3509
 Db 12610 GATGAAGCACCCCGTAGCCGACTTTGACGCAACAGGACCAACCCAGAAAGACAGATTT 12669
 QY 3510 ATTGACAGCCAGGACAGTAAAGCCACTGGAGCTGTGAGCAGTCAAGGTCAAGAGAA 3569
 Db 12670 GTGSAATTCATGAGGAGAAATGAGTCCGATCTGGTATCCCAAGTCAAGGGGACAGAAAGCA 12729
 QY 3570 GAGGCGGTACTGCTCAGAAAGAGGAGCTTCGACACTACCTTAATAATGTTCCAGCCAG 3629
 Db 12730 GAGGAGTTCCTGCAAGAAAGAGAGGCTTCAGCACCTTCCAGTTTGTGTTCAGGAA 12789
 QY 3630 GAAGAACATGGGGAAGAACCCAGGAAGAGATGTTCTTGAACCTTACACAGCAAGCTTACT 3689
 Db 12790 GAACTAAAGAACATCAAGATGGAAGACACTCTAGAGCATACAGATAAGAGGTGCA 12849
 QY 3690 GCTGAGCGGTGCCCGTCTTGGCAAGAGACTGAGGTGGGTCAAGAGGGTGTAG----- 3740
 Db 12850 GTGGAACCTGTATCCATTTCTGTCAAAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCT 12909

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:13:30 ; Search time 1074 Seconds
(without alignments)
12916.489 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 CGGCGCGGCGGAGTGAAG.....aataaacatttgaactac 6160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 437078

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_101002.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
25: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5182.4	84.1	5200	18 AAV02302	Rat tumour suppress
2	4696.6	76.2	5074	18 AAV02301	Rat tumour suppress
3	2179.6	35.4	6614	24 ABNS59861	Novel human coding
4	2144.4	34.8	6608	24 ABNS96942	Gene #3440 used to
5	2144.4	34.8	6608	24 ABL61938	Colon adenocarcino
6	2118.8	34.4	6605	19 AAV23545	Human gravin polyp
7	2118.8	34.4	6605	21 AAV74903	Human gravin codin
8	1920.6	31.2	6886	23 AAS95205	DNA encoding novel
9	550	8.9	563	24 ABK63316	Rat sequence diffe

10	172.2	2.8	377	20	AA040328	Human secreted pro
11	162.2	2.6	178	24	ABK63568	EST sequence diffe
12	105.2	1.7	483	20	AAV86620	EST clone BF146.
13	101.8	1.7	459	23	ABV56830	Human prostate exp
14	95.4	1.5	400	22	AAI86887	Human polynucleoti
15	92	1.5	339	20	AAV86330	EST clone AJ169.
16	81.8	1.3	10732	21	AAAI0594	Gene encoding a su
17	79.4	1.3	385	22	AAI83958	Human polynucleoti
18	78.6	1.3	3399	17	AAV05868	Chicken leucocytoz
19	76	1.2	49999	20	AAZ23891	Murine LOBO genom
20	76	1.2	49999	20	AAZ23896	Murine LOBO homolo
21	72.8	1.2	2800	22	AAZ07131	Canine retinitis p
22	72.4	1.2	1886	16	AAQ87587	DNA encoding leuco
23	72.2	1.2	6741	21	AAAI0595	Gene encoding a su
24	71	1.2	575	22	ABA50472	Human breast cell
25	71	1.2	575	22	ABA68422	Human foetal liver
26	71	1.2	575	22	ABA35413	Probe #13879 for g
27	71	1.2	575	22	AAK16793	Human brain expres
28	71	1.2	575	22	AAK42567	Human bone marrow
29	71	1.2	575	22	AAI23315	Probe #13248 for g
30	71	1.2	575	22	AAI48636	Probe #17322 used
31	71	1.2	575	22	AAI08956	Probe #8947 used t
32	71	1.2	575	24	ABS16615	Human genome-deriv
33	71	1.2	1969	22	ABA45341	Human breast cell
34	71	1.2	1969	22	ABA55830	Human foetal liver
35	71	1.2	1969	22	ABA25506	Probe #3972 for ge
36	71	1.2	1969	22	AAK04048	Human brain expres
37	71	1.2	1969	22	AAK29533	Human bone marrow
38	71	1.2	1969	22	AAI14105	Probe #4038 for ge
39	71	1.2	1969	22	AAI35486	Probe #4172 used t
40	71	1.2	1969	22	AAI03958	Probe #3949 used t
41	71	1.2	1969	24	ABS04084	Human genome-deriv
42	70.8	1.1	799	19	AAV55831	Nucleotide sequenc
43	70.8	1.1	1926	21	AAAS0254	Epstein Barr virus
44	70.8	1.1	1926	22	AAV82902	EBV tethering prot
45	70.8	1.1	2580	21	AAV75454	Nucleotide sequenc

ALIGNMENTS

RESULT 1
AAV02302
ID AAV02302 standard; cDNA; 5200 BP.
AC
XX AAV02302;
XX
XX
DT 20-JUL-1998 (first entry)
XX
XX Rat tumour suppressor gene SSeCKs.
DE
DE SSeCKs; tumour suppressor gene; rat; protein kinase C; mitosis;
KW cancer; malignancy; cell proliferation; Alzheimer's disease;
KW therapy; ss.
XX
OS Rattus sp.
XX
XX Key Location/Qualifiers
FT CDS 1..4791
FT /*tag= a
XX
XX WO9740059-A1.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-US06830.
XX
XX 18-JUN-1996; 96US-0665401.
PR 19-APR-1996; 96US-0635121.
XX
XX (GELM/) GELMAN I.
PA (JAKE/) JAKEN S.
XX

PI Gelman I, Jaken S;
 XX WPI, 1997-535770/49.
 DR P-PSDB; AAM31347.
 XX
 PT Tumour suppressor gene SSeCKs - used as a mitotic regulator, and
 PT inhibitor of malignant phenotype
 XX
 PS Claim 1; Fig 11A-L; 162pp; English.
 XX
 CC This rat cDNA sequence codes for the full-length form of the novel
 CC tumour suppressor protein SSeCKs (see AAM31347). The SSeCKs gene
 CC product is a substrate of protein kinase C and acts as a negative
 CC regulator of mitosis and as an inhibitor of the transformed
 CC phenotype. A partial SSeCKs clone (see AAV02301) was identified by
 CC searching for cDNAs whose abundance was low in NIH 3T3 cells and
 CC decreased following the expression of the activated oncogene v-src.
 CC A 5'RACE product was spliced to this truncated sequence to construct
 CC the full-length SSeCKs cDNA sequence. This full-length SSeCKs
 CC nucleic acid sequence, as well as homologous and hybridising nucleic
 CC acids are claimed, as are vectors comprising such nucleic acids,
 CC encoded proteins, host cells and methods of inhibiting the
 CC expression of a transformed phenotype in a host cell by introducing
 CC the nucleic acid. Introduction of a SSeCKs nucleic acid or gene
 CC product into a host cell inhibits mitosis of the host cell,
 CC allowing the treatment of diseases associated with disorders of
 CC proliferation and/or with the expression of a malignant phenotype.
 CC SSeCKs can also be used to treat or identify disorders of
 CC cytoskeletal structure and cellular architecture (such as
 CC Alzheimer's disease), and may be a marker for aberrancies in
 CC fertility and/or nervous system development.
 XX
 XX Sequence 5200 BP; 1612 A; 1201 C; 1563 G; 824 T; 0 other;
 Query Match 84.1%; Score 5182.4; DB 18; Length 5200;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 5189; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY CTCACTGTCAAGAAAGATGAAAGGGGAGAGAGCCCTCTGTGGAGCTGGAGACAC 632
 Db CTCACTGTCAAGAAAGATGAAAGGGGAGAGAGCCCTCTGTGGAGCTGGAGACAC 600
 QY CAGAGAGCCAGTGTGAGACTGCGTCGAGAGTCACATCTCAAGAAAGATGAGCTGAAG 692
 Db CAGAGAGCCAGTGTGAGACTGCGTCGAGAGTCACATCTCAAGAAAGATGAGCTGAAG 660
 QY CAATCCAAGAGAGCAAGAAAGGACCCCTGAAGCAAGACAGACACAGAAATCCCC 752
 Db CAATCCAAGAGAGCAAGAAAGGACCCCTGAAGCAAGACAGACACAGAAATCCCC 720
 QY CTTGAACCCGAAATCTGTCAAGCGGCTGAGAGAAAGCCCAAGATGAAAGAAAGAAAA 812
 Db CTTGAACCCGAAATCTGTCAAGCGGCTGAGAGAAAGCCCAAGATGAAAGAAAGAAAA 780
 QY CAAGAGAAAGAGCCCAAGATCCCGAGAGCCCAAGATCAACAGTGAACA 872
 Db CAAGAGAAAGAGCCCAAGATCCCGAGAGCCCAAGATCAACAGTGAACA 840
 QY ACATCTTCTTCAAGAAAGTCTTCACTACAGGTTGGCCGCTGGCCCAAGAACAGC 932
 Db ACATCTTCTTCAAGAAAGTCTTCACTACAGGTTGGCCGCTGGCCCAAGAACAGC 900
 QY TTCAAGAAATCTAAAGAGATGATCTGAGAACTGCGCAAGAAAGAGCAAGAGCA 960
 Db TTCAAGAAATCTAAAGAGATGATCTGAGAACTGCGCAAGAAAGAGCAAGAGCA 932
 QY GAAAAAGTGAAGAGAGAAAGAAAGAAAGCAAGAGCAGCTTGAGAGAGCAGAGCCG 1052
 Db GAAAAAGTGAAGAGAGAAAGAAAGAAAGCAAGAGCAGCTTGAGAGAGCAGAGCCG 1020
 QY GCAGAGAAACACAGACCAAGCCAGGTTGTACAGACATACAGAAAGTGAAGCTGCTTG 1112
 Db GCAGAGAAACACAGACCAAGCCAGGTTGTGTACAGACATACAGAAAGTGAAGCTGCTTG 1080
 QY GAAACACAGGTTGTGTACCTGAGAGCATCGTCAAGAGAAAGTGTGCTCTTTGGCAAG 1172
 Db GAAACACAGGTTGTGTGTACCTGAGAGCATCGTCAAGAGAAAGTGTGCTCTTTGGCAAG 1140
 QY GAAAGTTGTATGAGAAAGATGAAAGCCCAAGAAAGTGTGTGAGAGGTCACAGTAGC 1232
 Db GAAAGTTGTATGAGAAAGATGAAAGCCCAAGAAAGTGTGTGAGAGGTCACAGTAGC 1200
 QY GAAAGTTGTATGAGAAAGATGAAAGCCCAAGAAAGTGTGTGAGAGGTCACAGTAGC 1260
 Db GAAAGTTGTATGAGAAAGATGAAAGCCCAAGAAAGTGTGTGAGAGGTCACAGTAGC 1230
 QY ACCGTGAGAAAGACAGAGAGAGCAGAGAGAGAGAGAGCTGAAGAGAGGCTGCTG 1292
 Db ACCGTGAGAAAGACAGAGAGAGCAGAGAGAGAGAGAGCTGAAGAGAGGCTGCTG 1260
 QY GTAGAAGAAACAGAGAAATCTTTCCTCCCTGAGAAATCTGCTGAGAGCCCAAGAGTCCC 1352
 Db GTAGAAGAAACAGAGAAATCTTTCCTCCCTGAGAAATCTGCTGAGAGCCCAAGAGTCCC 1320
 QY GTAGAAGAAACAGAGAAATCTTTCCTCCCTGAGAAATCTGCTGAGAGCCCAAGAGTCCC 1380
 Db GTAGAAGAAACAGAGAAATCTTTCCTCCCTGAGAAATCTGCTGAGAGCCCAAGAGTCCC 1350
 QY CAGAGAGCTGAGCTGCTGAGAGAGCTGATGAAGAGAGAGAGTGTGTCTGAGAGGA 1412
 Db CAGAGAGCTGAGCTGCTGAGAGAGCTGATGAAGAGAGAGAGTGTGTCTGAGAGGA 1380
 QY GACCAACTCAATGACAGACCTTAAGTCTTGAAGAGAGAGCTGCCCAACACCCAGAA 1472
 Db GACCAACTCAATGACAGACCTTAAGTCTTGAAGAGAGAGCTGCCCAACACCCAGAA 1440
 QY GGCATTTGCTAGTGAAGTGAAGTGTCTGCTCTCAAGAAAGATCAAGAGTCAAGAGT 1532
 Db GGCATTTGCTAGTGAAGTGAAGTGTCTGCTCTCAAGAAAGATCAAGAGTCAAGAGT 1500
 QY CCTTGAAGAAATCTTCAAGAGCTCAAGAGCTTGAAGAGCTGTGGAGAGAAACAGAG 1592
 Db CCTTGAAGAAATCTTCAAGAGCTCAAGAGCTTGAAGAGCTGTGGAGAGAAACAGAG 1560
 QY GCGAAACAGAGAGTGGGGAGAGCAAGAGAGCTGAGAAATCAACACATTCACACCGAA 1652
 Db GCGAAACAGAGAGTGGGGAGAGCAAGAGAGCTGAGAAATCAACACATTCACACCGAA 1620

QY 1653 TCCCCAGAGAGTCTGATGACGACGAAGGAGAGAGCTCTGCGTCTGCTCCCGAGAGCGCT 1712
DB 1621 TCCCCAGAGAGTCTGATGACGACGAAGGAGAGAGCTCTGCGTCTGCTCCCGAGAGCGCT 1680
QY 1713 GAGGAGACCACTCTCTGGAGAAAGGCGCTGGAAGCACCCACAGGATGGGGAAGCTGAG 1772
DB 1681 GAGGAGACCACTCTCTGGAGAAAGGCGCTGGAAGCACCCACAGGATGGGGAAGCTGAG 1740
QY 1773 GAAGGAATCTACTCCGATGGAGAGAAGAAGAGAGAGGAGTCACTCCCTGGGCATCTTTC 1832
DB 1741 GAAGGAATCTACTCCGATGGAGAGAAGAAGAGAGAGGAGTCACTCCCTGGGCATCTTTC 1800
QY 1833 AAAAGATGGTGACACCCAGAAAGCGGTCCGAAGACCTTCTGAGATGACAAAGAGGAA 1892
DB 1801 AAAAGATGGTGACACCCAGAAAGCGGTCCGAAGACCTTCTGAGATGACAAAGAGGAA 1860
QY 1893 GAGCTGGAGAGGTCGAAGGCGCCACCTTGTCTCCACTGATAGCACAGTGTCAAGAAATG 1952
DB 1861 GAGCTGGAGAGGTCGAAGGCGCCACCTTGTCTCCACTGATAGCACAGTGTCAAGAAATG 1920
QY 1953 CAAGATGAAGTCAAAACTGTTGGTGGAGAAACAAAGCCAGAGGAACCAAGCGTAGGGTG 2012
DB 1921 CAAGATGAAGTCAAAACTGTTGGTGGAGAAACAAAGCCAGAGGAACCAAGCGTAGGGTG 1980
QY 2013 GATACCTTCACTGCTTGGGAAGCACTGATTTGTGTGCGATCATCAAGAGAGAGCAAGG 2072
DB 1981 GATACCTTCACTGCTTGGGAAGCACTGATTTGTGTGCGATCATCAAGAGAGAGCAAGG 2040
QY 2073 AAGGCATCTCTTCAGATGATGAAGGAGGCGCAAGGACACTGGGAGGGGACAGTCAAGAA 2132
DB 2041 AAGGCATCTCTTCAGATGATGAAGGAGGCGCAAGGACACTGGGAGGGGACAGTCAAGAA 2100
QY 2133 GCAGAGAGGCGCAAGAAAGAAAGAGCGGGAACAGACGCTGTTCTTCGCGAGACCCAG 2192
DB 2101 GCAGAGAGGCGCAAGAAAGAAAGAGCGGGAACAGACGCTGTTCTTCGCGAGACCCAG 2160
QY 2193 GAGCAGACCAAGCGCAAGGAGTTCTCTACCCAGCGAGCGGGAAGCGCTTCGGAAGGG 2252
DB 2161 GAGCAGACCAAGCGCAAGGAGTTCTCTACCCAGCGAGCGGGAAGCGCTTCGGAAGGG 2220
QY 2253 GAAGGTGTCTCCACTTGGGAGTCAATTTAAAGATTAGTCACTCCAAGAAATAATCCAAG 2312
DB 2221 GAAGGTGTCTCCACTTGGGAGTCAATTTAAAGATTAGTCACTCCAAGAAATAATCCAAG 2280
QY 2313 TCRAAATCGAAGAGAAAGCCGGAAGACTCTAGTGTAGAGCAGTGTCTCCACTGAGATCGAA 2372
DB 2281 TCRAAATCGAAGAGAAAGCCGGAAGACTCTAGTGTAGAGCAGTGTCTCCACTGAGATCGAA 2340
QY 2373 CCGAGTAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATPCCCGGACGCGGAAGAAA 2432
DB 2341 CCGAGTAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATPCCCGGACGCGGAAGAAA 2400
QY 2433 AGGSCAGACGGGAAGCAAGAACCACTGTGGAAGACTCAGGGCCAGTGGAGATAAT 2492
DB 2401 AGGSCAGACGGGAAGCAAGAACCACTGTGGAAGACTCAGGGCCAGTGGAGATAAT 2460
QY 2493 GAGGACGACCCCTAATGTCCAGCGCTGCTGCTCTGTCTGAGTATAATGCAAGTGGAGAGG 2552
DB 2461 GAGGACGACCCCTAATGTCCAGCGCTGCTGCTCTGTCTGAGTATAATGCAAGTGGAGAGG 2520
QY 2553 GAGAAGATGGNAGCCCAAGGGAATACGGAGTGCCCGAGCTGCTGGGGGCTGTGTAAGTG 2612
DB 2521 GAGAAGATGGNAGCCCAAGGGAATACGGAGTGCCCGAGCTGCTGGGGGCTGTGTAAGTG 2580
QY 2613 TCCGAGGAGCTCAGTAAGACTCTGCTCCACTGTGAGTGTCCGAGTCAATTTGAGGAGCC 2672
DB 2581 TCCGAGGAGCTCAGTAAGACTCTGCTCCACTGTGAGTGTCCGAGTCAATTTGAGGAGCC 2640
QY 2673 AGGSCAGTCAACAGTGTCAAGAGCGGTCTCCTTTCGTGGATATCCGCTTCGTTAAACAGAA 2732
DB 2641 AGGSCAGTCAACAGTGTCAAGAGCGGTCTCCTTTCGTGGATATCCGCTTCGTTAAACAGAA 2700
QY 2733 CCTCTTGAAACACACAGCGGGAGAGCCATGCCCTGTGTGAAGAGGTCACTGAAAAAGAC 2792

DB 2701 CCTCTTGAAACACACAGCGGAGAGCCATGCCACTCTGTTGAAGAGGTCACTGAAAAAGAC 2760
QY 2793 ATCAATTCAGAAAGAACTCTCTGTGCTCACCCAGACGTTTACAGAGGGTAAAGATGCCCAT 2852
DB 2761 ATCAATTCAGAAAGAACTCTCTGTGCTCACCCAGACGTTTACAGAGGGTAAAGATGCCCAT 2820
QY 2853 GACGACATGTCTCAACAGTGAAGTGGATTTTCACTTCAGAAAGCTGTGACAGCCACAGAGACC 2912
DB 2821 GACGACATGTCTCAACAGTGAAGTGGATTTTCACTTCAGAAAGCTGTGACAGCCACAGAGACC 2880
QY 2913 TCAGAGGCTCTCCGTACTCTGAGAGTGTACCGAAGCATCCGGGGCCCGAAGAGACACAGAGAC 2972
DB 2881 TCAGAGGCTCTCCGTACTCTGAGAAAGTGTACCGAAGCATCCGGGGCCCGAAGAGACACAGAGAC 2940
QY 2973 ATGCTGTCCGACGTTTCCAGCTGACTCTCCCGACACACACAGAGGAAGCCACCCCA 3032
DB 2941 ATGCTGTCCGACGTTTCCAGCTGACTCTCCCGACACACACAGAGGAAGCCACCCCA 3000
QY 3033 GTTCAGGAGTAGAGAGTGTGTCTAGATACAGAAAGAGGAGCGCCACAGAGCGAGGCC 3092
DB 3001 GTTCAGGAGTAGAGAGTGTGTCTAGATACAGAAAGAGGAGCGCCACAGAGCGAGGCC 3060
QY 3093 ATCTTCCAAAGCGTTTGCAGACAAAGTGAAGAGAGTCCCGAGTGCCTGCAACCCAGACT 3152
DB 3061 ATCTTCCAAAGCGTTTGCAGACAAAGTGAAGAGAGTCCCGAGTGCCTGCAACCCAGACT 3120
QY 3153 GTGCAGAGAAACGGGTCAAAAGCACTGGAGAAAGTGTGAGGAGGTAGAGGAGGACTCCGAA 3212
DB 3121 GTGCAGAGAAACGGGTCAAAAGCACTGGAGAAAGTGTGAGGAGGTAGAGGAGGACTCCGAA 3180
QY 3213 GTCTGTGCTTCCGAGAAAGAGAGGACGTTATGCCGAAAGGAGCCCGTGAGGAGGACTGGA 3272
DB 3181 GTCTGTGCTTCCGAGAAAGAGAGGACGTTATGCCGAAAGGAGCCCGTGAGGAGGACTGGA 3240
QY 3273 GCTGAGGACTTTCGACAGGCTCTGAGACTGAGACGCTGAGCAGGCTACTCCAGAGAGCCTTGAAGTT 3332
DB 3241 GCTGAGGACTTTCGACAGGCTCTGAGACTGAGCAGGCTACTCCAGAGAGCCTTGAAGTT 3300
QY 3333 CTTGAAGTCAAGCAGATGTAGACCATGTGCCAGCTGCGAGGTTATCAAGCTCCAGCAG 3392
DB 3301 CTTGAAGTCAAGCAGATGTAGACCATGTGCCAGCTGCGAGGTTATCAAGCTCCAGCAG 3360
QY 3393 CTGATGSAACAGGCGCGTGGCCCTCTGAGTCTATCCGAAACCTTTCAGACAGTGTGACACAAAT 3452
DB 3361 CTGATGSAACAGGCGCGTGGCCCTCTGAGTCTATCCGAAACCTTTCAGACAGTGTGACACAAAT 3420
QY 3453 GGAAGCACTCCCTTAGCAGATTCAGACACTGCGAGATGGGACACAGCAAGATGAAACCAAT 3512
DB 3421 GGAAGCACTCCCTTAGCAGATTCAGACACTGCGAGATGGGACACAGCAAGATGAAACCAAT 3480
QY 3513 GACAGCCAGGACAGTAAGCCACTGCGAGCTGTCCAGGAGTCCAGGTCACAGAGGAAGAG 3572
DB 3481 GACAGCCAGGACAGTAAGCCACTGCGAGCTGTCCAGGAGTCCAGGTCACAGAGGAAGAG 3540
QY 3573 CGGCTACTCTCTCAGAAAGAGGAGCCTTCCACACTACTTAATAATGTTCCAGGCCAGGAA 3632
DB 3541 CGGCTACTCTCTCAGAAAGAGGAGCCTTCCACACTACTTAATAATGTTCCAGGCCAGGAA 3600
QY 3633 GAACTGCGGGAAGAACCCAGGAAGAGATGTTCTTGAACCTTACAGCAAGAGTGTACTGCT 3692
DB 3601 GAACTGCGGGAAGAACCCAGGAAGAGATGTTCTTGAACCTTACAGCAAGAGTGTACTGCT 3660
QY 3693 GCAGCCGTGCGCTTCTGCGAAGAGACTGAGGTGGGTCAAGAGGTTGAGTGTGCTGTTG 3752
DB 3661 GCAGCCGTGCGCTTCTGCGAAGAGACTGAGGTGGGTCAAGAGGTTGAGTGTGCTGTTG 3720
QY 3753 GATGGAAAAAGTCAAAAGAAACAGGAGGTGTTGTACACTCTGACCCCAACAGTCAA 3812
DB 3721 GATGGAAAAAGTCAAAAGAAACAGGAGGTGTTGTACACTCTGACCCCAACAGTCAA 3780
QY 3813 AAGGCTGCTGATGTGACATATGACGTGAAGTGTGGAAGTGTGCGGTTGTGAGGAAAG 3872

```

Db 3781 AAGCTGCTGATGTGACATATGACAGTGAAGTATGGAGTGGCCGGGTCTCAGAAAAG 3840
Qy 3873 GAGAGTACTGAAAGTGCAGAGTCTTAAGCTTGAAGAGAGAGAGATGAAAATCTGACTTGA 3932
Db 3841 GAGAGTACTGAAAGTGCAGAGTCTTAAGCTTGAAGAGAGAGAGATGAAAATCTGACTTGA 3900
Qy 3993 AAGAGAAAAGAGAGACAAAGCCAGACAGTGAAGTGAAGAGTGAAGAGAGAGAGAGAGAGAG 3992
Db 3901 AAGAGAAAAGAGAGACAAAGCCAGACAGTGAAGTGAAGAGTGAAGAGAGAGAGAGAGAGAG 3960
Qy 3993 GCTCCTGAGCATGAAGAACTTACCGGAAAGCCAGTCTGACACTTGATGCGCAGCTCA 4052
Db 3961 GCTCCTGAGCATGAAGAACTTACCGGAAAGCCAGTCTGACACTTGATGCGCAGCTCA 4020
Qy 4053 GAGAGGGGAAAGGCACTGGGAAAGCCTTGGAGAGAGCCCTTCTCTCCAGACCAAGACAA 4112
Db 4021 GAGAGGGGAAAGGCACTGGGAAAGCCTTGGAGAGAGCCCTTCTCTCCAGACCAAGACAA 4080
Qy 4113 GCAGGTTCATGAGATGTTCAAGTTCAAAGCTTGACACAAAGTCACTCAACACAGAGA 4172
Db 4081 GCAGGTTCATGAGATGTTCAAGTTCAAAGCTTGACACAAAGTCACTCAACACAGAGA 4140
Qy 4173 GCTGTGAAAAAGGTCATGAAACCGTGTGATTTGATTTGAGAGACAGTGAAGTCCAGAGTGT 4232
Db 4141 GCTGTGAAAAAGGTCATGAAACCGTGTGATTTGATTTGAGAGACAGTGAAGTCCAGAGTGT 4200
Qy 4233 GTAGGTGACACTTATTAACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGACTCTT 4292
Db 4201 GTAGGTGACACTTATTAACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGACTCTT 4260
Qy 4293 CAGCATGCAAGAGACACGGTACCCCTGGGGCTGTAGTCTCAGAGCAATTCATCCCAATC 4352
Db 4261 CAGCATGCAAGAGACACGGTACCCCTGGGGCTGTAGTCTCAGAGCAATTCATCCCAATC 4320
Qy 4353 ATAGTAATCTCTGCTCTCTGAAAGAACCCCTACATCTCTGACCTTCAAGAGAAATTAAGCGA 4412
Db 4321 ATAGTAATCTCTGCTCTCTGAAAGAACCCCTACATCTCTGACCTTCAAGAGAAATTAAGCGA 4380
Qy 4413 TCCCAAGAGAGACGATCAGAGAGAGAGACAAAGCAGATGCTGTGCTCTGATGCTGACGGC 4472
Db 4381 TCCCAAGAGAGACGATCAGAGAGAGAGACAAAGCAGATGCTGTGCTCTGATGCTGACGGC 4440
Qy 4473 AAGGAGATGACGCAATGCAAAAAGTCTCAAGGCTGAACCTGAATCTCTGAACTTGAAG 4532
Db 4441 AAGGAGATGACGCAATGCAAAAAGTCTCAAGGCTGAACCTGAATCTCTGAACTTGAAG 4500
Qy 4533 AGTAAGAGCAACAAAGTGTGTGTAACGTCATTCAGACAGCGGTTGACAGTTCGACAGT 4592
Db 4501 AGTAAGAGCAACAAAGTGTGTGTAACGTCATTCAGACAGCGGTTGACAGTTCGACAGT 4560
Qy 4593 ACAGAAACAGCCCCCGAAAATCATGTGTTATGATTCACAGACCCAGGTTCTGATGACAG 4652
Db 4561 ACAGAAACAGCCCCCGAAAATCATGTGTTATGATTCACAGACCCAGGTTCTGATGACAG 4620
Qy 4653 CTGGACAGAGAGAGCCCAACAGATGCTGACAAATAAAGATGCCAAGATGAAGACAC 4712
Db 4621 CTGGACAGAGAGAGCCCAACAGATGCTGACAAATAAAGATGCCAAGATGAAGACAC 4680
Qy 4713 CCAGTCCGACAGCCCAAGAGAGACTTGAAGTCTGACAGTCTGAGAGCATGAGGCTCAG 4772
Db 4681 CCAGTCCGACAGCCCAAGAGAGACTTGAAGTCTGACAGTCTGAGAGCATGAGGCTCAG 4740
Qy 4773 CCTCGAAAATGCTTCCCGGCTTGGAGTTGAAAGCGCCGGTCTCAAAGTAAAGCATTTAGA 4832
Db 4741 CCTCGAAAATGCTTCCCGGCTTGGAGTTGAAAGCGCCGGTCTCAAAGTAAAGCATTTAGA 4800
Qy 4833 AGCTGCTCTCTCAACCCAAAGATCCAAAAGAGCATGCTGATGAGCCCTCAGTCCAA 4892
Db 4801 AGCTGCTCTCTCAACCCAAAGATCCAAAAGAGCATGCTGATGAGCCCTCAGTCCAA 4860
Qy 4893 AGCTTAAGCCCAAGAGAGAGCCAGTCTGAGAACTTAACCAAGAAATCCCAAGACAC 4952
Db 4861 AGCTTAAGCCCAAGAGAGAGCCAGTCTGAGAACTTAACCAAGAAATCCCAAGACAC 4920

```

```

Qy 4953 AACGACCAAAAGCTTAACCGAGAGAGGCGGATCCCCAAAAGTTGAGGTCCAGAAAGAA 5012
Db 4921 ACCGAGCCAAAGCTTAACCGAGAGAGGCGGATCCCCAAAAGTTGAGGTCCAGAAAGAA 4980
Qy 5013 ATGTCTACCAAGTCAGTCAAAAGAGAACAGGCCAGGACAGAGAGACCTGACAGAGCA 5072
Db 4981 ATGTCTACCAAGTCAGTCAAAAGAGAACAGGCCAGGACAGAGAGACCTGACAGAGCA 5040
Qy 5073 AAGGAGACCTGGCAGAAATCCTTAAGATGTTAGTGTCTCATTTGATCATCTGTAAGACCA 5132
Db 5041 AAGGAGACCTGGCAGAAATCCTCGATGTTAGTGTCTCATTTGATCATCTGTAAGACCA 5100
Qy 5133 ATGTGAAAACAAAGTCACAGAAACAGATGCTGTGTGGACCTTGAGACCAAGATTTTCA 5192
Db 5101 ATGTGAAAACAAAGTCACAGAAACAGATGCTGTGTGGACCTTGAGACCAAGATTTTCA 5160
Qy 5193 AGCCCATGAGATCCAGAGAGAGGCGGTCCTCAATGATTTT 5232
Db 5161 AGCCCATGAGATCCAGAGAGAGGCGGTCCTCAATGATTTT 5200

RESULT 2
AAV02301
ID AAV02301 standard; cDNA; 5074 BP.
XX
AC AAV02301;
XX
DT 20-JUL-1998 (first entry)
XX
DE Rat tumour suppressor gene SSeCKS (truncated form).
XX
KW SSeCKS; tumour suppressor gene; rat; protein kinase C; mitosis;
KW cancer; malignancy; cell proliferation; Alzheimer's disease;
KW therapy; ss.
XX
OS Rattus sp.
XX
FH Key
FH CDS
FT 176..4216
FT /*tag= a
FT /transl_except= (pos:389..391, aa:Glu)
FT /transl_except= (pos:1661..1663, aa:lys)
FT /transl_except= (pos:1668..1670, aa:lys)
FT /transl_except= (pos:2444..2446, aa:Arg)
FT /transl_except= (pos:2528..2530, aa:Glu)
FT 2528..2530
FT /*tag= b
FT polyA_signal
FT 5056..5061
FT /*tag= c
FT
XX
PN W09740059-A1.
XX
PD 30-OCT-1997.
XX
PE 18-APR-1997; 97MO-US060830.
XX
PR 18-JUN-1996; 96US-0665401.
XX
PR 19-APR-1996; 96US-0635121.
XX
PA (GELM/) GELMAN I.
PA (JAKEN/) JAKEN S.
XX
PI Gelman I, Jaken S;
XX
DR WPI: 1997-535770/49.
DR P-PSDB: AAM31346.
XX
PT Tumour suppressor gene SSeCKS - used as a mitotic regulator, and
XX inhibitor of malignant phenotype
XX
PS Example 6; Fig 3A-H; 162pp; English.
XX

```

CC This rat cDNA sequence codes for an active truncated form of the
CC novel tumour suppressor protein SseCKs (see AAW31346). The SseCKs
CC gene product is a substrate of protein kinase C and acts as a
CC negative regulator of mitosis and as an inhibitor of the transformed
CC phenotype. The SseCKs clone was identified by searching for cDNAs
CC whose abundance was low in NIH 3T3 cells and decreased following the
CC expression of the activated oncogene v-src. A 5' RACE product was
CC spliced to the truncated sequence to construct a full-length
CC SseCKs cDNA (see AAV02302). This full-length SseCKs nucleic acid
CC sequence, as well as homologous and hybridising nucleic acids, are
CC claimed, as are vectors comprising such nucleic acids, host cells
CC and methods of inhibiting the expression of a transformed phenotype
CC in a host cell by introducing such nucleic acids. Introduction of
CC a SseCKs nucleic acid or gene product into a host cell inhibits
CC mitosis of the host cell, allowing the treatment of diseases
CC associated with disorders of proliferation and/or with the
CC expression of a malignant phenotype. SseCKs can also be used to
CC treat or identify disorders of cytoskeletal structure and cellular
CC architecture (such as Alzheimer's disease), and may be a marker
CC for aberrancies in fertility and/or nervous system development.
XX
SQ Sequence 5074 BP; 1501 A; 1151 C; 1441 G; 981 T; 0 other;

Query Match 76.2%; Score 4696.6; DB 18; Length 5074;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 4985; Conservative 0; Mismatches 79; Indels 91; Gaps 15;

QY 1016 GGAAGAAGACAGAGCCAGCTCGAGGAGCAGAGCCGGCAGACACACAGACAGGCGCAG 1075
DB 1 GGAAGAAGACAGAGCCAGCTCGAGGAGCAGAGCCGGCAGACACACAGACAGGCGCAG 60

QY 1076 GTTGTACGACGACTACAGAGAGTGGAGCTGCCTTTTGGAGACACAGGTTGGTACCTGGA 1135
DB 61 GTTGTACGACGACTACAGAGAGTGGAGCTGCCTTTTGGAGACACAGGTTGGTACCTGGA 120

QY 1136 GGCATCTCAGAGAGAGTGTCTCTTTGGCAACCGAAGTGTTCATGAGAAGATGGA 1195
DB 121 GGCATCTCAGAGAGAGTGTCTCTTTGGCAACCGAAGTGTTCATGAGAAGATGGA 180

QY 1196 AGCCCAACAAAGAGTGTGTGACAGAGTCCACGTGAGACACCGTGGAGAGACAGAGGAGGA 1255
DB 181 AGCCCAACAAAGAGTGTGTGACAGAGTCCACGTGAGACACCGTGGAGAGACAGAGGAGGA 240

QY 1256 GCAGGAGGAGGAGGAGGAGGCTGAAGGGGGCGTGGTGGTGAAGAAACAGAGAAATCCTT 1315
DB 241 GCAGGAGGAGGAGGAGGAGGCTGAAGGGGGCGTGGTGGTGAAGAAACAGAGAAATCCTT 300

QY 1316 GCCCCCTGAGAACTGCTGAGCCCCCAGAGAGTCCCCCAGGAAGCTGAGCCTGCTGAGGA 1375
DB 301 GCCCCCTGAGAACTGCTGAGCCCCCAGAGAGTCCCCCAGGAAGCTGAGCCTGCTGAGGA 360

QY 1376 GCTGATGAAGACAGAGAGATGTGTCTCTGGAGGAGACACACTCAACTGACAGACCT 1435
DB 361 GCTGATGAAGACAGAGAGATGTGTCTCTGGAGGAGACACACTCAACTGACAGACCT 420

QY 1436 AAGTCTCTGAAGAGAGAGCTGCCCAACACCCAGAGAGGATTCAGTGAAGTGGAGAT 1495
DB 421 AAGTCTCTGAAGAGAGAGCTGCCCAACACCCAGAGAGGATTCAGTGAAGTGGAGAT 480

QY 1496 GCTGTCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTTGAAGAAATCTTTCACTAG 1555
DB 481 GCTGTCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTTGAAGAAATCTTTCACTAG 540

QY 1556 CTCAGGCTTAAAGAGAGTGTCTGGAGAGACAGAGGGGAACAGAGAGTGGGGGAGA 1615
DB 541 CTCAGGCTTAAAGAGAGTGTCTGGAGAGAGACAGAGGGGAACAGAGAGTGGGGGAGA 600

QY 1616 CGAAGAGCTGGAGATACCAACACATTTCACACCGAATCCCCAGAGAGTGTCTGATGAGCA 1675
DB 601 CGAAGAGCTGGAGATACCAACACATTTCACACCGAATCCCCAGAGAGTGTCTGATGAGCA 660

QY 1676 GAAGGGAGAGAGCTCTCGCTGCTCCCCGAGAGGACCTTGAGAGACACACGCTGTCTGGAGAA 1735
DB 1735 GAAGGGAGAGAGCTCTCGCTGCTCCCCGAGAGGACCTTGAGAGACACACGCTGTCTGGAGAA 1787

DB 661 GAAGGGAGAGAGCTCTCGCTGCTGCTCCCCGAGAGGACCTTGAGGAGACACCGTGTCTGGAGAA 720

QY 1736 AGGGCCGCTGGAAGACACCCAGAGATGGGAAGCTGAGGAGGAATCTACTTCCGATGGAGA 1795

DB 721 AGGGCCGCTGGAAGCA-CCAGAGTGGGAGCTGAGGAGGAATCTACTTTC-----GTGGA 775

QY 1796 GAAGAAGAGAAAGAGGATCACTCCCTGGGCATCTCTTCAAAAAGATGGTGAACCCCAAGAA 1855

DB 776 GAGAAGAAGAGGAAGGATCACTCCCTGGGCATCTCTTCAAAAAGATGGTGAACCCCAAGAA 835

QY 1856 ACCGGTCCGAAGACCTTCTGAGAGTGCACAGGAGGAGAGCTGGAGAAGTCAAGAGCGC 1915

DB 836 AC--GGTCCGAAGACCTTCTGAGAGTGCACAGGAGGAGAGCTGGAGAAGTCAAGAGCGC 894

QY 1916 CACCTTGTCTCCACTGATAGCACAGTGTGAGAAATGCAAGATGAAATCAAAATCTGTTGG 1975

DB 895 CACCTTGTCTCCACTGATAGCACAGTGTGAGAAATGCAAGATGAAATCAAAATCTGTTGG 954

QY 1976 TGAGGAACAAAAGCCAGAGAACCAAGCGTAGGGTGGATCTTTCAGTGTCTTGGGAAGC 2035

DB 955 TGAGGAACAAAAGCCAGAGAACCAAGCGTAGGGTGGATCTTTCAGTGTCTTGGGAAGC 1014

QY 2036 ACTGATTTGTGTCGGATCATCTCAAGAGAGAGCAAGCAAGGCGATCCTCTTCAGATGATGA 2095

DB 1015 ACTGATTTGTGTCGGATCATCTCAAGAGAGAGCAAGCAAGGCGATCCTCTTCAT---GATAT 1071

QY 2096 AGGAGGCCCAAGGACACTGGGA--GGGGACAGTCAAGAGCAGAGGAGGCGCAGCAAGACA 2154

DB 1072 AAGAGGCCCAAGGACACTGGGAGGGGACAGTCAAGAGCAGAGGAGGCGCAGCAAGACA 1131

QY 2155 AAGAAGCCGGAACAGACGCTGTTCTGCCAGACCCAGGAGCAGGACCAAGCGCAAGGAA 2214

DB 1132 AAGAAGCC--GAACAGACGCTGTTCTGCCAGACCCAGGAGCAGGACCAAGCGCAAGGAA 1190

QY 2215 GTTCTCTACCCGAGCAGCGGGAAGCCCTTCCGAAGGGGAGAGTGTCTCCACTTGGGAGT 2274

DB 1191 GTTCTCTACCCGAGCAGCGGGAAGCCCTTCCGAAGGGGAGAGTGTCTCCACTTGGGAGT 1250

QY 2275 CATTTAAAAGATTAGTCACTCCCAAGAAAAAATCCAAAGTCAAAATCTGGAAGAGAAAGCCG 2334

DB 1251 CATTTAAAAGATTAGTCACTCCCAAGAAAAAATCCAAAGTCAAAATCTGGAAGAGAAAG 1310

QY 2335 AAGACTCTAGTGTAGACA-----GTTGTCACTGAGATCGAACCCGAGTAGAGAAGA 2386

DB 1311 CCGGAAGGACTCTAGTGTAGGAGCAGGTTGTCCACTGAGATCGAACCCGTTAGAGAAGA 1370

QY 2387 ATCTTGGTGTTCATTAAAGAAATTCATCCCGACCGCGGAAGAAAGGCGAGACGGGAA 2446

DB 1371 ATCTTGGTGTTCATTAAAGAAATTCATCCCGACCGCGGAAGAAAGGCGAGATGGGAA 1430

QY 2447 -GCAAGAACAAAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGACGACCCCTA 2505

DB 1431 GGCAAGAACAAAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGACGAGCCTG 1490

QY 2506 ATGTCCCAGCCGTCGTGCCCTCTGTCTGAGTATAATGAGTGGAGGAGGAAGATGGAAG 2565

DB 1491 ATGTCCCAGCAGTCGTGCCCTCTGTCTGAGTATGATGACGTGGAGGAGGAAGATGGAAG 1550

QY 2566 CCAGGGGGAATACGAGAGTGTGCCCGACCGTGTGGGGGCTGTGACGTGTCCGAGAGCTCA 2625

DB 1551 CCAGGGGGAATGCGGAGTGTGCCCGACCGTGTGGGGGCTGTGTA---GTGTCCGAGAGCTCA 1607

QY 2626 GTAAGACTCTGTGTCACACTGTGAGTGTGCGAGTCAATGATGGGACCCAGGGCAGTCA 2685

DB 1608 GTAAGACTCTGTGTCACACTGTGAGTGTGCGAGTCAATGATGGGACCCAGGGCAGTCA 1667

QY 2686 GTGTGGAAGAGCGGTCTCTTTCGTGGATATCCGCTTCCGTAACAGAACTCTTTGAACACA 2745

DB 1668 GTGTGGAAGAGCGGTCTCTTTCGTGGATATCCGCTTCCGTAACAGAACTCTTTGAACACA 1727

QY 2746 CAGCGGAGAGCCATCCACCTGTTGAGAGAGTCACTGAAAAGACATCATTTGCGAGAG 2805

DB 1728 CAGCGGAGAGCCATCCACCTGTTGAGAGAGTCACTGAAAAGACATCATTTGCGAGAG 1787

QY 2806 AAACCTCTGTGCTCAACCCAGAGCTTTACAGAGGGTAAAGATGCTCATGACCATGTGTCA 2865
DB 1788 AAATCTCTGTGCTCAACCCAGAGCTTTACAGAGGGTAAAGATGCTCATGACCATGTGTCA 1847
QY 2866 CCAGGAAAGTGGATTTTCACTCCAGAGGCTGTGACAGCCACAGAGACTCCAGAGGCTCTCC 2925
DB 1848 CCAGGAAAGTGGATTTTCACTCCAGAGGCTGTGACAGCCACAGAGACTCCAGAGGCTCTCC 1907
QY 2926 GTACTGAAGAATTACCGAAGCATCGGGGCCGAAGAACAACAGACATGTGTCCGACG 2985
DB 1908 GTACTGAAGAATTACCGAAGCATCGGGGCCGAAGAACAACAGACATGTGTCCGACG 1967
QY 2986 TTTTCCAGCTGATGATCTCCAGACACCAACAGAGAAAGCCCACTTCAGAGGATG 3045
DB 1968 TTTTCCAGCTGATGATCTCCAGACACCAACAGAGAAAGCCCACTTCAGAGGATG 2027
QY 3046 AGAGTGTGTGTAGATTCAGAGAAGAGAGAGCCGACAGAGGCACTCCCAAGCCG 3105
DB 2028 AGAGTGTGTGTAGATTCAGAGAAGAGAGAGCCGACAGAGGCACTCCCAAGCCG 2087
QY 3106 TTGCAGACAAGGTGAAGAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGACAGAAACG 3165
DB 2088 TTGCAGACAAGGTGAAGAGAGAGTCCAGAGTGCCTGCAACCCAGACTGTGACAGAAACG 2147
QY 3166 GGTCAAAAGCACTGGAGAGAGGTGAGAGAGTGAAGAGAGACTCCGAAGTGTGCTTCG 3225
DB 2148 GGTCAAAAGCACTGGAGAGAGGTGAGAGAGTGAAGAGAGACTCCGAAGTGTGCTTCG 2207
QY 3226 AGAAGAGAAGAGACCTTATGCGGAAGAGCCGCTGACAGAGACTGAGAGTGAACCTTTG 3285
DB 2208 AGAAGAGAAGAGACCTTATGCGGAAGAGCCGCTGAGAGAGACTGAGAGTGAACCTTTG 2267
QY 3286 CACAGGGCTCTGAGACTGAGACAGGCTACTCAGAGAGCCCTGAAGTTCCTGAAGTCAAG 3345
DB 2268 CACAGGGCTCTGAGACTGAGACAGGCTACTCAGAGAGCCCTGAAGTTCCTGAAGTCAAG 2327
QY 3346 CAGATGTAGACATGTCCGACAGTGCAGGTTATCAACTCCAGACATGATGAAACAG 3405
DB 2328 CAGATGTAGACATGTCCGACAGTGCAGGTTATCAACTCCAGACATGATGAAACAG 2387
QY 3406 CCGTGGCCCTGAGTCACTCCGAACCTTGAACAGACAGTGAACAATGGAAACATCCCT 3465
DB 2388 CCGTGGCCCTGAGTCACTCCGAACCTTGAACAGACAGTGAACAATGGAAACATCCCT 2447
QY 3466 TAGCAGATTCAGACACTGAGATGGGACACAGCAAGATGAAACATTTGACAGCAGACA 3525
DB 2448 TAGCAGATTCAGACACTGAGATGGGACACAGCAAGATGAAACATTTGACAGCAGACA 2507
QY 3526 GTAAAGCCACTGACAGCTGTCAAGGCAAGTCAAGAGTCAAGAAAGAGGCGCTACTGCTC 3585
DB 2508 GTAAAGCCACTGACAGCTGTCAAGGCAAGTCAAGAGTCAAGAAAGAGGCGCTACTGCTC 2567
QY 3586 AGAAAGAGAGGCTTGCACACTTAATTAATGTTCCAGCCAGGAAGAACTGGGGAAG 3645
DB 2568 AGAAAGAGAGGCTTGCACACTTAATTAATGTTCCAGCCAGGAAGAACTGGGGAAG 2627
QY 3646 AACCGAAGAGATGTTCTTGAACCTTACAGCAAGAGCTTACTGTGACGCGTCCCG 3705
DB 2628 AACCGAAGAGATGTTCTTGAACCTTACAGCAAGAGCTTACTGTGACGCGTCCCG 2687
QY 3706 TTCTGGCAAAAGACTGAGGTGGGTCAAGAGGTTGAGTGTGATGAGAAAG 3765
DB 2688 TTCTGGCAAAAGACTGAGGTGGGTCAAGAGGTTGAGTGTGATGAGAAAG 2747
QY 3766 TCAAGAGAAGAAAGAGAGGTGTTGTACACTGTGACCCCAAGCTCAAAAGGCGCTGATG 3825
DB 2748 TCAAGAGAAGAAAGAGAGGTGTTGTACACTGTGACCCCAAGCTCAAAAGGCGCTGATG 2807
QY 3826 TGACATATGACAGTGAAGTGAAGTGGAGTGGCGGCTGTGAGAGAAAGAGATGCTGAAG 3885
DB 2808 TGACATATGACAGTGAAGTGAAGTGGAGTGGCGGCTGTGAGAGAAAGAGATGCTGAAG 2867

QY 3886 TGCAAGTCTTAGCTTGAGAGGAGAGATGGAACGTGAAAGAGAGAAAGG 3945
DB 2868 TGCAAGTCTTAGCTTGAGAGGAGAGAGATGGAACGTGAAAGAGAGAAAGG 2927
QY 3946 AGACAAAGCCAGAGCAAGTGAAGTGAAGAGTGAAGAGAGAAACAGCCGCTCTGAGATG 4005
DB 2928 AGACAAAGCCAGAGCAAGTGAAGTGAAGAGTGAAGAGAGAGAAACAGCCGCTCTGAGATG 2987
QY 4006 AAGAACCTTACGGAGAGCACTGACACTTGAACATGCCCAGCTCAGAGAGGGGAAG 4065
DB 2988 AAGAACCTTACGGAGAGCACTGACACTTGAACATGCCCAGCTCAGAGAGGGGAAG 3047
QY 4066 CACTGGAGAGCTTGAAGAGAGCCCTTCTCCAGACCAAGACAAAGACAGGTTGACATG 4125
DB 3048 CACTGGAGAGCTTGAAGAGAGCCCTTCTCCAGACCAAGACAAAGACAGGTTGACATG 3107
QY 4126 AGGTTCAAGTTCAAAAGCTTGACACAACAGTCACTCAACAGCAGAAAGCTGTGAAAAAG 4185
DB 3108 AGGTTCAAGTTCAAAAGCTTGACACAACAGTCACTCAACAGCAGAAAGCTGTGAAAAAG 3167
QY 4186 TCATPAAACGGTTGTGATTTCAAGACAGGTGAAGTCCAGAGTGTGATGTCACACT 4245
DB 3168 TCATPAAACGGTTGTGATTTCAAGACAGGTGAAGTCCAGAGTGTGATGTCACACT 3227
QY 4246 TATTACCAAGCTGAAGAGCTCTGCAACGGGTGCACTGGAACCTTCAGAGATGCAAGG 4305
DB 3228 TATTACCAAGCTGAAGAGCTCTGCAACGGGTGCACTGGAACCTTCAGAGATGCAAGG 3287
QY 4306 ACAGGTAACCCCTGGGGGCTGAGTCTCAGCAGAAATCCATCCCAATGATTAATCTCCTG 4365
DB 3288 ACAGGTAACCCCTGGGGGCTGAGTCTCAGCAGAAATCCATCCCAATGATTAATCTCCTG 3347
QY 4366 CTCTGAAAGCACTTACATCTCTGACCTTCAAGAGAAATTAAGGCAATCCCAAGAGAGC 4425
DB 3348 CTCTGAAAGCACTTACATCTCTGACCTTCAAGAGAAATTAAGGCAATCCCAAGAGAGC 3407
QY 4426 GATCAGAGAAGAGAGCAAGCCAGATGCTGCTGATGCTGACAGGCAAGAGAGATGAG 4485
DB 3408 GATCAGAGAAGAGAGCAAGCCAGATGCTGCTGATGCTGACAGGCAAGAGAGATGAG 3467
QY 4486 CAATGAAAAAGTCTCTCAAGGCTGAACCTGAGATCTGGAACCTTGAAGTGAAGACA 4545
DB 3468 CAATGAAAAAGTCTCTCAAGGCTGAACCTGAGATCTGGAACCTTGAAGTGAAGACA 3527
QY 4546 AGATTGTCTGAAGCTTATGACAGAGCGTTGACAGCTTGGACAGTACAGAAACAGCC 4605
DB 3528 AGATTGTCTGAAGCTTATGACAGAGCGTTGACAGCTTGGACAGTACAGAAACAGCC 3587
QY 4606 CCGAAACTCATGCTTATGATTCAGACAGACCCAGGTTCCGTCATGACAGGTTGACAGAGG 4665
DB 3588 CCGAAACTCATGCTTATGATTCAGACAGACCCAGGTTCCGTCATGACAGGTTGACAGAGG 3647
QY 4666 AGCCCAACAGATGCTGACCAAAATGAAGATGCAAGATGAAACACCCATGCGCAGC 4725
DB 3648 AGCCCAACAGATGCTGACCAAAATGAAGATGCAAGATGAAACACCCATGCGCAGC 3707
QY 4726 CCAAGAGAGACTTGCAGTCTGACCGTTCGAGAGCAATGAGGCTTCAAGCTTCAAGAACT 4785
DB 3708 CCAAGAGAGACTTGCAGTCTGACCGTTCGAGAGCAATGAGGCTTCAAGCTTCAAGAACT 3765
QY 4786 TGCCCGGCTTGCAGTGAAGCCGCGGTGCAAGTGAAGTGAAGAGTGCCTCTCA 4845
DB 3766 TGCCCGGCTTGCAGTGAAGCCGCGGTGCAAGTGAAGTGAAGAGTGCCTCTCA 3825
QY 4846 ACCCAAGATCCAAAGAGAGATGCTGATGGCCCTCACTCCAAAGCTTGAAGCCAGG 4905
DB 3826 ACCCAAGATCCAAAGAGAGATGCTGATGGCCCTCACTCCAAAGCTTGAAGCCAGG 3884
QY 4906 CAGAGGCAAGTCTTGAAGAACTTAACCAAGAAATCCCAAGACCAACGAGCAAGG 4965
DB 3885 CAGAGGCAAGTCTTGAAGAACTTAACCAAGAAATCCCAAGACCAACGAGCAAGG 3941
QY 4966 TAAACGAGAGAGGCAATCCCAAGAAAGTTGAAGTCCAGAGAAAGAAATGCTTACCAAGT 5025

|||||
Db 3942 TAA CGGAGGCG-ATGCCCCAAAAGTTGAGTTCAGGAAGAAATGCTACCAAGT 4000
Qy 5026 CAGTCAAAGAGAAACAAGGCCCGCAGGAGAGGACCTGCGAGGAGCCAAAGGAGGACCTGG 5085
Db 4001 CAGTCAAAGAGAAACAAGGCCCGCAGGAGAGGACCTGCGAGGAGCCAAAGGAGGACCTGG 4060
Qy 5086 CAGATCCTAAGATGTTAGTTGCTCATTTGATACATCTGTAAAGACCAAGATGTGAACAAAG 5145
Db 4061 CAGATCCTAAGATGTTAGTTGCTCATTTGATACATCTGTAAAGACCAAGATGTGAACAAAG 4120
Qy 5146 TCACAGAACAAAGATGCTGCTGCGGACCTTGAGACCAAGATTTTCAGAGCCCATGAGATC 5205
Db 4121 TCACAGAACAAAGATGCTGCTGCGGACCTTG-GACCAAGATTTTCAGAGCCCATGAGATC 4179
Qy 5206 CAGAGACAGGGCGCTCCAATGATTTCCACCAGTAGAGCACCCCGACAAATTCAGAGGCT 5265
Db 4180 CAGAGACAGGGCGCTCCAATGATTTCCACCAGTAGAGCACCCCGACAAATTCAGAGGCT 4239
Qy 5266 TCATCGGAGCTAGAGCAGCTAAATTCCTCGTTTCAAGACTGCGCTTGAATTTGGCCCC 5325
Db 4240 TCATCGGAGCTAGAGCAG----- 4259
Qy 5326 TTGATGCGTCCGCTGATTTCTCAACATTTCTCGTTTCAAGACTGCGCTTGAATTTGGCCCC 5385
Db 4260 -----CTAACATTTCTCGTTTCAAGACTGCGCTTGAATTTGGCCCC 4299
Qy 5386 TTGATGCGTCCGCTGATTTTCGGATTTAAGGTCTCGGTTCTCAACCTTGAACCAATTC 5445
Db 4300 TTGATGCGTCCGCTGATTTTCGGATTTAAGGTCTCGGTTCTCAACCTTGAACCAATTC 4359
Qy 5446 GCCATACCTAGTTCACCTTCAAACTGGAGCATCTCTTTATGATTTATATATGATG 5505
Db 4360 GCCATACCTAGTTCACCTTCAAACTGGAGCATCTCTTTATGATTTATATATGATG 4419
Qy 5506 TTTATGAGTCTCTCTGACCTATGATATATTTTCTAACGTTAAGCATGCT 5565
Db 4420 TTTATGAGTCTCTCTGACCTATGATATATTTTCTAACGTTAAGCATGCT 4479
Qy 5566 TTTTGTATTATGCAATATATAACGGGTGTGAGCATAGCAGCTTTGAAAAGCTCCAA 5625
Db 4480 TTTTGTATTATGCAATATATAACGGGTGTGAGCATAGCAGCTTTGAAAAGCTCCAA 4539
Qy 5626 GCCTCAACTGTAACTGCAGCAACAGATAAATTCCTGCGCAAGAGAGACAAGTCTTT 5685
Db 4540 GCCTCAACTGTAACTGCAGCAACAGATAAATTCCTGCGCAAGAGAGACAAGTCTTT 4599
Qy 5686 TTAAGTTTACTGATGCTTAGATCTGGGCTTCTAGTCTCTGAAAGTGGTTTTC 5745
Db 4600 TTAAGTTTACTGATGCTTAGATCTGGGCTTCTAGTCTCTGAAAGTGGTTTTC 4659
Qy 5746 TATGCACAGCGAGCTCAGAAATAAAACCCCATTTTGAAACATCCAGGATGTCCTCAATAT 5805
Db 4660 TATGCACAGCGAGCTCAGAAATAAAACCCCATTTTGAAACATCCAGGATGTCCTCAATAT 4719
Qy 5806 TACCATGATTTTCCCGCTTTTGTGTAATCCAGTCCAGTTGGAAAGAGTCTCCTCT 5865
Db 4720 TACCATGATTTTCCCGCTTTTGTGTAATCCAGTCCAGTTGGAAAGAGTCTCCTCT 4779
Qy 5866 GTGTCAGATTAAAGCCCTGCTCTTAATGATATGACAAATGAGTGTGCTTAAGCCATGA 5925
Db 4780 GTGTCAGATTAAAGCCCTGCTCTCTTAATGATATGATGGACAAATGAGTGTGCTTAAGCCATGA 4839
Qy 5926 GATGTTTCTTAATGCAGAGGAATCTGTTGACGTTTTTTTGTGTTGATGTTCTCTATGCT 5985
Db 4840 GATGTTTCTTAATGCAGAGGAATCTGTTGACGTTTTTTTGTGTTGATGTTCTCTATGCT 4899
Qy 5986 GGACCGAATTCATATGCAGATCGAAGTGAGTCTCTGTTCTTTTACAGATGGTATTTGATAG 6045
Db 4900 GGACCGAATTCATATGCAGATCGAAGTGAGTCTCTGTTCTTTTACAGATGGTATTTGATAG 4959
Qy 6046 ATACTGAGTTTCTGTGTTATATCTGTGCCCCCTTTTAAAGAACAAATGTGCAATTATG 6105
|||||

Db 4960 ATACTGAGTTTCTGTGTTATATCTGTGCCCTTCTTTAAGAACAAATGTGCAATTATG 5019
Qy 6106 TTCCTTTGGATAAATTTGATTTGACAACTGATTTAAATAAACAATTTTCACTAC 6160
Db 5020 TTCCTTTGGATAAATTTGATTTGACAACTGATTTAAATAAACAATTTTCACTAC 5074
RESULT 3
ABN59861
ID ABN59861 standard; cDNA; 6614 BP.
XX ABN59861;
AC
XX
XX 28-JUN-2002 (first entry)
XX
XX Novel human coding sequence SEQ ID NO: 272.
XX Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX Homo sapiens.
OS
XX
XX WO200222660-A2.
PN
XX
XX 21-MAR-2002.
PD
XX
XX 10-SEP-2001; 2001WO-US26015.
PF
XX
XX 11-SEP-2000; 2000US-0659671.
PR
XX
XX (HYSE-) HYSEQ INC.
PA
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-292408/33.
DR P-PSDB; ABB97448.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
PS Claim 1; SEQ ID NO 272; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
XX invention.
SQ Sequence 6614 BP; 2094 A; 1420 C; 1827 G; 1273 T; 0 other;
Query Match 35.4%; Score 2179.6; DB 24; Length 6614;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 4223; Conservative 0; Mismatches 1794; Indels 537; Gaps 43;
Qy 3 GGCCGGCGGGAGTAGAGGCCACTGAGCCATGGGCGCAGGAGTTCACCGAGCAGCG 62
Db 162 GGCTAGCGCGGAGAGATGCGGAGAGCCATGGGCGCGGAGTCCACCGAGCAGCG 221
Qy 63 AGCCCGAGCAGCAGCG--GCGGGAGCGACACGCCGAGGAGCTGTCTCAGTGGCCAT 119
Db 222 AGCCCGAGCAGCAGCGCGCGGAGGAGCTCCACGCCGCTGAGCCCGAGCCAGCGGCG 281
Qy 120 GGCCCGCAGCTGAAGCCTCGGA-----GCAGCTGGAGACCCCGCGGCGGAC 170
|||||

Db	282	GGCCCTCGGCGGAGGGGGCCAGACACACC	GGGAGACCCCGCCATGCTGCTCGGAC	341
Oy	171	CCGCGCCACCAAGCTCCCAAGAAAGAA	TGGCCAGCTGTTCTGTCAACGCGTGA	230
Db	342	CCGCGCCACCAAGCTCCCAAGAAAGAA	TGGTCACTGTCACATCAATGGGTAGTGA	401
Oy	231	CAAGAGATGTCAATGTTCCAAAGGAAA	CCAGAGGG-----	268
Db	402	CAGATAGCTCAGCCTCCGAGGGGTGAC	CTPMAATGGCCAGAAAGAGCCCTGAACGTT	461
Oy	269	-----	GCAGAGGAAAGATCGTTGATGAGAA	308
Db	462	CAAGAGCCTTAAACAGCCAGAGGAAAG	AAAGTCAATTGTCAAGAGGTTGGACAGAG	521
Oy	309	GAGTCAGAAAGATGTGAGAGAAAAAG	CCAGTGTGAAGAAATGGCGCCACCTCCACAGT	368
Db	522	GACTCTGAAGATGTGAGCGAAAGAG	CTCGATPAAAGATGGCTACTTAATGACGGTT	581
Oy	369	GTTGAGATATTACAAAAGATTGGCGAG	GAGAGACATCAGAAATTAATTGAACAAATCCCT	428
Db	582	GTTTACACACATCACAGATGTGGCGAG	GAGAGACACCCGAATTAATGAACAATTCCT	641
Oy	429	GCTTCAGAAAACAATGTGGAAGAAAG	TGTAACGCTGCTGAGTCCAGGCTPAATGATT	488
Db	642	TCTTCAGAAAGCAATTTAGAAAGCT	PAACAAACCACATGATCCAGAGCTPAATGATT	701
Oy	489	GGCTTCAGAAAGATTAATTAATTTG	TTGTTTAAATTCACGCTGAAGAGATPAAAT	548
Db	702	GGATTTAAGAAAGTGTTAAGTTTGT	TGGCTTTAAATTCATGTGAAGAAAGATPAGCA	761
Oy	549	GAAGAATCAATACTGTCCAATCTCA	CTCTCAGAGAGATGAAGCGGAAGGGCGAA	608
Db	762	GAGAAAGCTGACCTGTCCAGCTACT	CACTGTGAAGAAAGATGAAGGGGAGAGACGA	821
Oy	609	GCCTCTGTGCGAGCTGGAGACCA	CCAGAGGCCAGTGTGAACACTGCCGTGGAGATCA	668
Db	822	-----	GGGGTGGCGACACACAGAGCCCAAGCCTTGGGGCT-----	GGAGAGCA 866
Oy	669	GCATCCAAAGAAATGTAGCTGAAG	CAATCCACAGAGACGAAGAGCAACCTTGAAGCA	728
Db	867	GCATCCAAAGAAAGCCGAACCCAA	ACAATCTACAGAGAAACCGAAGAGCCCTGAAGGT	926
Oy	729	GAACAGACAGACAGAAATCCCCTT	CAAGCCGAATCTGATCAACGGCTGAGAGAGAA	788
Db	927	GAGCAAAAGCCAGCAAAATTTCT	CCCCCAGCCGAATCTGGCGAAGCA---	GTGAGAGAA 983
Oy	789	GCCAAAGATGAAGAGAAAGAAACA	AGAAAGAACCCCAAGTCCCGAATTCGCCG	848
Db	984	TGCAAAAGAGAAAGAGAAAGAAACA	AGAAAGAACTTAGCAAGTCTGCAGAAATCTTCG	1043
Oy	849	AGCAGCCCAATCACTGTGAGACAC	ACTCTTCCTTCAAGAGTTCTTCACTCACGGTTGG	908
Db	1044	ACTAGTCCCGTGAACAGTGAACAG	ATCACTTGAAGAAATTTCTTCACTCAAGTTGG	1103
Oy	909	GCCGCTGGCGCAAGAAAGACCA	GCCTTCAAGAAATCAAAAGAGATGATCTGGAAC	CTGCC 968
Db	1104	GCCGCTGGCGCAAAAGAACCA	GTTTCAGAAAGCCAGAAAGATGAAGTGAAGCTTCA	1163
Oy	969	GAGAAAGAAAGAGCAAGAGCGCA	AGAAAAAGTGAACAGAGAAAGAAAGAAAGACAGAG	1028
Db	1164	GAGAAAGAAAGAAAGCAAGAGCC	AGAAAAAGTGAACAGAGAAAGAAAGACAGAG	1223
Oy	1029	CCAGCTGTGGAGGA-----	GCAGAGCGCGCAGAGAAC	1061
Db	1224	GTTGCTTCGAGAAACTGACCG	CTCCGAGCAAGCCACACAGAGCCGGCAGAAAGT	1283
Oy	1062	ACAGACCAAGCCAGGTTGTACG	ACAGACTTACAGAGAGTGAAGCTGCTTTGGAAGCAG	1121
Db	1284	GCCCAAGAGCCCGGTTATACAG	CTGAATATGAGAAAGTTGAGCTGCCCTCAGAGAGCA	1343
Oy	1122	GTTGTGACTTGAGGCAATCGT	CAGAGAGAAATGTGCTCTTTGGCAACGGAATGTTT	1381
Db	1344	GTCAGTGGCTGCGAGGGACCTT	CTGAAGAAACATGTGCTCGTTGGCGCAGAAATGTTT	1403

QY	1182	GAAAGAAATGGAACCCACCAAGAA---	GTGTGTGAGAGGCTCAAGTGAACCGTG	1238
Db	1404	GATGAGAAATAGAGTCCACCAAGAGAGTTGTGGCCGAAGTCACTGACACCGTG		1463
QY	1239	GAGAAGA CAGAGAGAGCAGGAGAGAGAGAGAGAGCTGAAGGGAGCGTGTGTGA		1298
Db	1464	GAGGAGGAACCGAAGCGAGAA-----	-ACGAGGTGAA	1499
QY	1239	GGAA CAGAGAAATCTTGGCCCCCTAGAAATCTGCTGAGCCCCAGAGATCCCCAGAA		1358
Db	1500	GAACACACAGGCTCTGTGGCAGCTGAAGATTGTGTGAATGATATGCAAACTCCAGAA		1559
QY	1359	GCTGAGCTGCTGAGGAGCTGATGAAGAGACAGAGATGTGTCTCTGAGAGAGACAC		1418
Db	1560	GCCAACTTCCCAAGAGAGCTGTGGTAAGCTCAAGAAAGATGTGTGTTCCGAGAGAGACCT		1619
QY	1419	ACTCACTGACAGACCTTAAGTCTTGAAAGAGACGCTGGCCAAACCCAGAAAGCATT		1478
Db	1620	ACAAGGAGAGCTGACCTCAGTCTCGATGGAAGGTGTGTCCAAACCCCGAAGGGCTT		1679
QY	1479	GTCAGTAGGAGAGATGCTGTCTCTCAGGAAATATCAAGATACAGGGAGTCCCTTG		1538
Db	1680	GTGAGTAGGTGAATGCTGTATATCACAGAGAAATGAAGGTGCAGGGAGATCCATA		1739
QY	1539	AAGAACTCTTCAGTAGCTCAGGCTTAAAGAGCTGTCTGGAGAGAGCAGAGGGGAA		1598
Db	1740	AAGAAAGCTTTTACAGCACTGGCTTAAAAAGCTTTCTGGAAAGAAACAAAGGGAA		1799
QY	1599	CGAGAGGTGGGGAGACGAAAGACCTGGAATATCAACACATTTACACCGAATCTCCCA		1658
Db	1800	AGAGGA---GGAGAGACGAGAAATAGGGGAGACACTCAGGTTCACCGAATCTTCG		1856
QY	1659	GAGAGTCTGATGAGCAAGAGAGAGAGACTCTGCGTGTCCCGCCGAGAGACCTGAGAG		1718
Db	1857	GACAGCCAGAGAGAGCAAAAGGGCCGAGACTCTGCTCATCTCCCTGAGAGACCCGAGAG		1916
QY	1719	ACCACGTCTGGAAGAAAGGGCCGCTGGAAGCACTCCAGATGGGGAGCTGAGAAAGA		1778
Db	1917	ATCAGTGTCTGAAAAAGGGCTTAGCCGAGGTGCAGACAGATGGGGAGCTGAAAGAA		1976
QY	1779	ACTATCTCCGATGAGAGAAAGAGAGAAAGGATACATCCCTGGGGCATCTTCAAAA		1838
Db	1977	GCTACTTCCGATGAGAGAAAAAAGAAAGGTGTCACTCTCTGGGCACTATTCAAAA		2036
QY	1839	ATGTGACACCCAAAGAAACGGGTCCGAAGACCTTCTGAGATGACAGAGAGAAAGCTG		1898
Db	2037	ATGTGAGCGCCCAAGAGACGTGTAGACGGCTTCGAAAGTGAATAAGATGAGCTG		2096
QY	1899	GAGAAAGTCAAGAGACGCCACCTTGTCTCTCACTGATAGCAAGTGTCAAGAAATCCAGAT		1958
Db	2097	GACAAAGTCAAGAGACCGCTACTCTGTCTTCCACCGAGACACGCTCTGAATCCAGAA		2156
QY	1959	GAAATCAAACTGTTGTGAGAAACAAAAGCCAGAGAAACCAACCTGAGGGTGAATACT		2018
Db	2157	GAAATGAAGAGGACGTGTGAGAGAGCCAAAGCCGAAAGAACAAACGCAAGTGTATCC		2216
QY	2019	TCAGTGTCTGGGAAGCACTGATTGTGTGATCATCCAGAAAGAGACAAAGAGCA		2078
Db	2217	TCAGTATCTTGGGAAGCTTTAATTTGTGTGGATTCATCCAAAGAAAGACAAAGGAAAG		2276
QY	2079	TCTCTTTCAATGATGAAGAGGGCCAAAGACATGGGAGGGGAACGTACAGAGCAGAG		2138
Db	2277	TCTCTTTGTGATGAGAAAGGGGACCAAAAGCATGTGGAGAGACACACGAAAGCTGAT		2336
QY	2139	GAGGCACCAAGACAAAGAAAGCCGGAACGAGAGCTGTTCTGCGACAGACCCAGAGCAG		2198
Db	2337	GAGGCCGGAAGAGCAAGAGACGGGGACAGAGGGGATCTCTTGCTGTGTTCCAGAAACAT		2396
QY	2199	GACCAAGCGCAAGAGTCTCTACCCGAGCCAGCGGGAGGCCCTTTCCGAGGGGAAGT		2258
Db	2397	GATTCAGGGCAGGGAGAGTTCTCTCCCGAGCAAGCTGGAAGGCCCTTACCGAAGGGAGGGC		2456

QY 2259 GTCTCCACTTGGAGTCATTTTAAAGATTAGTCACTCCAGAAAAAATCCAACTCAAAA 2318
Db 2457 GTTTCACCTCGGAGTCATTTAAAGGTTAGTCAGCCCAAGAAAAAATCAAAGTCAAG 2516
QY 2319 CTGGAAGAGAAAGCCGAGAC-----TCTAGTGTAGAGCAGTGTGTCACCTGAG 2366
Db 2517 CTGGAAGAGAAAGCCGAGACTCCATAGCTGGGTCTGTTGAGAACATTCACCTCCAGAC 2576
QY 2367 ATCGAAACCGAGTAGAAGAAATCTTTGGGTTTCCATTTAAGAAATTCATCCCGGAGCGGG 2426
Db 2577 ACTGAACCCGGTAAGAGAAATCTGGGTCTCAATCAAGAAATTTATTTCTTGGAGGAGG 2636
QY 2427 AAGAAAAAGGCGAGCGGAGCAAGAACAGCCACTGTGAAGACTCAGGCGCAGTGGAG 2486
Db 2637 AAGAAAAAGGCGAGTGGAAACAAGAACAGCCCTGTGTAAGACGCGAGGCGCAACAGGG 2696
QY 2487 ATAAATGAGGACGACCTTAATGTCCAGCCGTCGTCTGTCTGTAGTATAATGCAAGTG 2546
Db 2697 GCCAACGAATGACTCTGATGTCGCCGCCGTGTCCTCTGTCTGTAGTATGATGCTGTA 2756
QY 2547 GAGAGGAGAAAGATGG-----AAGCCAGGGGAATACGGAGCTGCCCCAGCTGCTG 2597
Db 2757 GAAAGGGAGAAAAATGGAGGCGACAGCAAGCCCAAAAAGCGCGAGAGCAGCCCGACGAAG 2816
QY 2598 GGGCTGTGTACGTCTCGAGGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTCCGA 2657
Db 2817 GCAGCCACTGAGGTGTCCAGGAGCTCAGCGAGAGTCAGGTTCATATGATGGCAGCAGCT 2876
QY 2658 GTCAATTGATGGGACCGGCGAGTCACCAAGTGTGAAAGCGGTCTCTCTGTGATATCC 2717
Db 2877 GTCGCTACGGGACGAGGCGAGTACCATTTATTTGAAAGAGTCTCTCTTTGATATCT 2936
QY 2718 GCTTCGTAAACAGAACTCTTTGAAACACACAGCGGGAGAGCCATGCCACCTGTTTGAAGAG 2777
Db 2937 GCTTCAGTGCAGAACTCTTTGAACAAGTAGAAGCTGAAGCCGCACTGTTAACTGAGGAG 2996
QY 2778 GTCACTGAAAGACATCATTCG---AGAAGAACTCTGTGTCTCACCGACGATTACCA 2834
Db 2997 GTATTGAAAGAGAAATTTGCAAGAAAGAAACCCCGGTTACTGAAACCTCTGCCA 3056
QY 2835 GAGGTTAAAGATGCCATGACGACATGTCACAGTGAAGTGAATTTCACTCAGAGAGCT 2894
Db 3057 GAGNACAGAGAGGCCCGGGGCGACACGGTGTGTTAGTGAAGCGGAATTTGACCCCGAAGCT 3116
QY 2895 GTGACGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAAGTTTACCGAAGCATCCGGG 2954
Db 3117 GTGACAGCTGCAGAAACTGCAGGGCCATTTGGTTTCCGAAGAAGAAACCGAAGCATCTGCT 3176
QY 2955 GCCGAAGAGACCAAGACATGTTGTCGCGAGTTCAGAGTTCACAGCTGACTGACTCCCGACACC 3014
Db 3177 GCTGGAAGAGACCAAGAAAATGGTGTGACAGCTCTCCAGTTAACCGACTCCCGACAGACC 3236
QY 3015 ACAGAGGAAGCCACCCAGTTTCAGAGGTGAGAGTGTGCTGTAGATACAGAGAGAGAG 3074
Db 3237 ACAGAGAGGCCATCTCCGTGCGAGAGGTGAAGTGGCGTACTGTACATAGAAGACAA 3296
QY 3075 GAGCGCCAGACGCGAGGCCATCTCCAGCCGTTGCGACAGAGGTGAAGAGGAGTCCCGAG 3134
Db 3297 GAGAGCGGACTCAAGAGTCTCCAGCAGTGCAGAAAAGTGAAGAGGAATCCAG 3356
QY 3135 GTGCTCTCAACC---CAGACTGTGAGAGAACCGGGGTCAAAAGCACTGGGAAGGTTGAG 3191
Db 3357 CTGCTGGCACCGGTGGGCGCAGAAAGATGTGCTTCAGGCTGTGCGAGAGCAGAGGCGAGAA 3416
QY 3192 GAGGTAGAGGAGGACTCCGAAGTCTCGCTTCGGAGAAAGAGAGGAGCTTATGCCAAA 3251
Db 3417 AGACCAGAGAGCGGCTGAAGCGTCCGGTCTGAAGAAAGAGAGCGGATGTAGTGTGAAA 3476
QY 3252 GGACCCGTGCGAGGAGCTGGAGCTAGCATCTTTGCACAGGGCTCTGAGACTGGACAGGCT 3311
Db 3477 GTAGATGCTCAGGAGGCAAAAATGAGCCTTTTACACAGGGAAGGTGGTGGGCGAGACC 3536
QY 3312 ACTCCAGAGAGCCCTTG---AAGTTCTGAAAGTCAAGCAGATGTAGAC-----CAT 3359

Db 3537 ACCCCAGAAAAGCTTTTGAAAAAGCTCTCTCAAGTCACAGAGAGCATAGAGTCCAGTGAGCTT 3596
QY 3360 GTGCGCACAGTGCCA-----GGTTATCAAGCTCCAGCAGAGCTGATGGAA 3401
Db 3597 GTAACCACTTGTCAAGCCGAAACCTTTAGCTGGGTAAATACACAGGAGATGGTATGGAA 3656
QY 3402 CAGGCGGTGGCCCTCGAGTCATCCGAAACCTTGCACAGACAGTGAAGCAAAATGGAAGCACT 3461
Db 3657 CAGGCTATCCCCCTGACTCGGTGGAAACCCCTACACACAGTGAAGTGTGATGGAAGCACC 3716
QY 3462 CCCTTAGCAGATTTCAGACACTGCGAGATGGGACACACAGACAGATGAACCAATTTGACAGCCAG 3521
Db 3717 CCCGTAGCCGACTTTGACGCAACCAGGCAACCCAGAAAGCAGAGATTGTGGAATCCAT 3776
QY 3522 GACAGTAAAGCCACTGTCAGCTGTGAGCAGTCAAGGTACAGAGTACAGAAAGAGGGCGGTACT 3581
Db 3777 GAGGAGNATGAGTGCATCTGTAACCACTGAGGGGGCACAGACAGAGGAGCAGTTCCT 3836
QY 3582 GCTCAGAAAGAGAGGCTTCGACACTACCTAATAATTTTCCAGCCCGAGGAAGAACATGGG 3641
Db 3837 GCACAGAAAGAGAGGCTTCAGCACCTTCAGTTCAGTTTGTGTTCCAGGAAGAACTAAAGAA 3896
QY 3642 GAAGAACCCAGGAAGAGATGTTTGAACCTACACAGCAAGAGCTTACTGTCAGCCGTG 3701
Db 3897 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTGAGTGAACCTGTA 3956
QY 3702 CCCGTTCTGCAAAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3957 TCCATTTCTGTCAAAGACTGAGGGGACTCAAGAGGCTGACCAAGTATGCTGATGAAAAACC 4016
QY 3753 GATGGAGAAAAAGTCAAGAAAGAACAGGAGGTGTTTGTACACTCTGG-----ACCCAAC 3806
Db 4017 AAAGACGTACCAATTTTTCGAAGGACTTGAGGGGTCTATAGACACAGCATACAGTCAGT 4076
QY 3807 AGTCAAAAGGCTGCTGATGTGACATATGACAGTGAAGTGAATGGAGTGGCGGGGTGTGAG 3866
Db 4077 CGGGAAGAGGTCTCACTGAAGTTGCCCTTAAAGGTGAAGGGAAGAAGAGTGAATGTAAA 4136
QY 3867 GAAAAGCAGAGTACTCAAGTGCAGAG-----TCTTAGCCTGGAGGAGGA 3911
Db 4137 AAGGATGATGCTCTTGAATGACAGAGTCAAGTCTCTCTCCATCCCGCTGGAGAGA 4196
QY 3912 GAGATGGAAAACCTGACGTTGAAAAGGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGTGA 3971
Db 4197 GAGATGTTAGTTTCAAGTTCGAAGGGGAGAAAACAGAAAGCAGAGCCAAACCCATGTGAATGA 4256
QY 3972 GAAGG---TGAGCAGGAAAACAGCCGCTCTGAGCATGAAGAAACCTACGGGAAGCCAGTC 4028
Db 4257 GAGAAGCTTGAGCACGAAAACAGCTGTTACCGTATCTGAAGAGGTTCAGTGAAGCAGCTCCTC 4316
QY 4029 CTGACACTTGCATGCCAGCTCAGAGAGGGGGAAGGCACTGGGAAGCCCTGGAGGAGAC 4088
Db 4317 CAGACAGTGAATGTGCCCATCATAGATGGGGCAAGAAAGTTCAGAGTTTGGAAAGAGAC 4376
QY 4089 CTTCTCTC---TCCAGACCAAGACAAAGAGGTTGATAGAGGTTCAAGTTTCAAAGCCTG 4145
Db 4377 CTTCTCTCTGCTAGGTCAAGAGGAGGAGTATGACCAAAATTCAGTTTCAAGTTCAGAGCTCT 4436
QY 4146 GACAAACAGTCTACTCAACAGCAGAAAGCTGTGGAANA-----GGTCATAGAAAACGGTT 4199
Db 4437 GAGGCATCATTTCACTCTAAACAGCGGCTGCAGAGGAGGAAAAGGTCTTAGGAGAAACTGCC 4496
QY 4200 GTGATTTTCAGAGACAGTGAAGTCCAGAGTGTGATAGGTGCACACTTATTACAGCTGAG 4259
Db 4497 AACATTTTAGAAAACAGGTGAACGTTGGAGCCTGCGAGGTCGACATTTAGTTCTTGGAAAG 4556
QY 4260 AAGTCTCTGCAACCGGTGGCCACTGGACTCTTTCAGCATGCGAGAGACAGCGTACCCCTG 4319
Db 4557 AAATCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAAGATGCTGTGCCACA 4616
QY 4320 GGGCTTGAGTCTCAGCGAGAAATCCATCCCAATCATAGTAACCTCTCTCTCTGTAAGACACC 4379

4617 GGGCCCACTGTCAGGCAAAATCCAGACAGTATGATCTGCTACTTACCAAGAAAGGC 4676
4380 CTACATCTCTGACTTACAGAGAGAAATTAAGCGCATCCAGAGAGAGCGATCAGAGAGAG 4439
4677 TTAAGTTCGACCTGAGAGAGAGAGAAACCAATCATGAAATGAGATGAGATGATGATC 4736
4440 GACAAAGCAGATGCTGCTGCTGATCTGACGCGAAGAGAGATGACGATGAGAAATGTC 4499
4737 GATGAGCAGGTGCTTGGC-----AGGAGGTCAAAGTGAATGATGAGAGAT 4790
4500 CTCAAGGCTGAACCTGAGATCTGAGACTTGAAGTAAGCAACAATTTGCTGAAC 4559
4791 TTAGAGCTGAAATGAGATTTTGAAGCTTGAAGCAAAAGAGATTTCTCCAAAC 4850
4560 GTCAATTCAGAGAGCGCTTGACCGATTCGACCGTAC---AGAAACAGCCCCGAGAACTCAT 4616
4881 ATCATCCAGACAGCGCTTGACCGATTTGATGAGTGAAGAAACAGCCAGCAATGTTG 4910
4617 GCTTATGATTCACAGACCGAGTTC-----CTGANTCGAGCTTGACAG 4660
4911 ACGTCTGAGTTACAGACCAAGCTCAGTATMAAGCTGACAGCGCAGAGCGCTGAGAG 4970
4661 CAGGAGCGCCACAGATGCTGAGCAAAATGAAAGATGCGAAGATGAAACACCCAGTGC 4720
4971 GAACCGAGAAAGAGAGAGAGAGACCTCAGGCTCTGACAGAGATGAAACCAATTAAT 5030
4721 GCAGCCCA-----GAGAGACTTGAAGTCTGACCGTCTGGA 4759
5031 TCAGCCAAAGAGAGATCAGAGTCAACCGCAGTGGGCAAGACATTTCTGATTTCCAAA 5090
4760 GGCATGGCTCAGCTCGGAAATGCTGCGCGCTTGACAGTTGAAAGCCCGGTGCA- 4818
5091 GACATGAGTGAAGCTCAGAAAGAACATGACTGTTGAGTGAAGTTCACTGTAAT 5150
4819 ----- 4818
5151 GATCAGCAGCTGGAAGAGTCTCTCCCATCTGAGAAAGAGAGTGAAGTGAACA 5210
4819 -----AGTAAGCATTTGAGAACTGCT 4840
5211 AAGTCTGTCAGAGATGATGTCATGCTTGTAGAGAAAGATAGAAAGTACATA 5270
4841 CCTCAACCCAAAGA--TCCAAAGAGACATGCTGCTGATGAGCCCTCAGCTCCAAAGCTTA 4898
5271 GTTGAACCGAAAGAGATGAAAGAGTGAATGTTGATGAGACCTGMAAACCAAGAACTCA 5330
4899 GCGCAGGAGAGAGCGAGTCTCTGGAACCTAACCAAGATCCCGACAGACCAAGCGA 4958
5331 GCGCTGGCTGATATGATGATGCTCAGAGAGCTTAACCAAGATCCCGACATACAAATGGA 5390
4959 CCAAGCTTAAACGAGAGGCGATCCCGCAAAAGTTGAGGTCAGAGAAAGAAATGCT 5018
5391 CCAAAACAAAGAGAGAGAGATGCGCAGAGAAATGTAATTCAGAGAAAGAAAGTGCAC 5450
5019 ACCAAGTCACTAAAGAGAAAGAGCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5078
5451 AGTGAATCAGATTAAGCGGATCAACCCCAAGCAAGAGAGAGATTAAGAAACAAAGGAA 5510
5079 GACCTGGCAGATTCCTAAGATGTTAGTTG-----CTCATTTGATC 5117
5511 GATCTGCAAGTCAAGATCTTACAGATCTTAAACATATGATGATTAACATTTGCT 5570
5118 ATCTGTAAGACAGATGTAAGAAACAGTCAAGAAACA-----GATGCTGCTTTGGAGC 5173
5571 GTTTGAAAGACAGATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5629
5174 CTTGAGACCAAGATTTCAAGAGCCCATGATCCAGAGAGAGAGAGAGAGAGAGAGAGAG 5233
5630 CTGAAGACCAAGATTTCAAG-ACCTTGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAG 5688
5234 ACCCAGTGAAGACACCCGCAATTTGAGAGCTTCACTGAGAGAGTGAAGCAGATTAAT 5293
5689 TTTTGAAGAGC-CCCTGAACATCTGAGGCTTCACTAGAGAGTGAAGCATTTAACATT 5747

5294 TCCCTGTTCAAGACCTGCTTGAATTTGCCCTTGATGAGCCGTCGCTATTTCTAACATT 5353
5748 TCCCTTTCCAAAGCAACCTACATATTTTCCCTTGATTAAC----- 5787
5354 TCCCTGTTCAAGACTGCTTGAATTTGCCCTTGATGAGCCGTCGCTATTTGAGATTTA 5413
5788 -----CATTAATAATTCGATTTA 5805
5414 AGTCTGCGCTTCAACCTGGAAC-CAATTTGCGCATTAACCTTAAGTTCACCTTCAAACT 5472
5806 AGTCTTAATTTCTTAACCTGGAACCTGAGTGGCAATTAACCTTAAGTTCGCTTGAAGT 5865
5473 GGAGATCTCTCTTAAGATTTATATGATGTTTATATGATGTTTATATGATGTTTATAT 5532
5866 GGAGATCAATCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5922
5533 TGTATATTTTCTTAACGTT-----AAGCATGCTTTTGTATTA 5575
5923 TGTATATTTTCTTATGTTTAAAGAAATGTCAGAGATTAATGATGTTTGTATCA 5982
5576 TGAATATATTAACGAGTGTGACGCAATAGCGAGCTTTGAAAGCTCCAAAGCTCAACG 5635
5983 CACAGTATATGATGGGAGATGCTCATATAGTGCAGAGCTTGGAGCTTTAAGCTCAGTTA 6042
5636 TAACCTGAGCAACATTAATTC-----CTGGCAAGAGAGACAAGCTTTT 5686
6043 TATTAACCAAAAG 6102
5687 TAAAGTTACTGATGCTTGAATCTGTGGGCTTCTAGTCTCTGAAAGTGTGTTTCT 5746
6103 TAAATTCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6182
5747 ATGCAACGAGCTCAGAAATTAAGAAACCCCAATTTGAAACATTCAGAGATGCCAATAT 5806
6163 ATTACAGAGAGTGTGCCAAACATAAAGCATTTTGAACATTAAGATGTTCTATGTC 6222
5807 ACCATGATTTTTCCTCCCTTTTCTGATATTCAGATCCAGAGTGGAGAAAGATCTCTCTG 5866
6223 ATTGGAAATTTTGC-----TTTCAACCCAGTGGAGGTTGAAGAGATTAATTTCT 6275
5867 TGTGAGA-----TTAAGCCCTGCTCTTAATATATGACAAATGAGT 5909
6276 GGTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6335
5910 GTGCTTAAGGCGATGAGA-----TGTTCCTAATGCAAGAGATCTGT 5953
6336 GTGCTTAATCCGAGGCAAGATGAAATATGTTTATATGTTAATGAAAGAAAGATGT 6395
5954 TGTAGTTTTTTTGAATGATCTCTTCTATGCTGAGCCGAATTCATATGAGATGGAAGT 6013
6396 TGTAG--TTTTTGAATCTTATATATGCTGAGCTGATTCACATGCGATGAATA 6453
6014 AGTCTGTTCTTACAGATGATTTGATATGATATGATATGATATGATATGATATGATATG 6073
6454 AGTCAAGTCTTACAAATGATTTTATATGATATGATATGATATGATATGATATGATATG 6513
6074 TGCCCTTCTTAAAGCAATGATGATTAATGTTCTTTGATTAATTTGATTTGACAA 6133
6514 TGCCATCTTTTAAAGCAATGATGCAACATTCATTTGATTAAGTTGATTTGACGA 6573
6134 CTGATTTAATTA 6147
6574 CTGATTTAATAAA 6587

RESULT 4
ABN96942
ID ABN96942 standard; DNA; 6608 BP.
XX
AC ABN96942;
XX
DT 13-AUG-2002 (first entry)

XX Gene #3440 used to diagnose liver cancer.
XX DE
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
XX Homo sapiens.
XX OS
XX W0200229103-A2.
XX PN
XX 11-APR-2002.
XX PD
XX 02-OCT-2001; 2001WO-US30589.
XX PF
XX 02-OCT-2000; 2000US-237054P.
XX PR
XX (GENE-) GENE LOGIC INC.
XX PA
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
PI WPI; 2002-426119/45.
XX DR
XX Diagnosing and detecting the progression of liver cancer.
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample
XX
XX Claim 1; SEQ ID NO 3440; 298pp; English.
XX PS
XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytotostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6608 BP; 2091 A; 1418 C; 1827 G; 1272 T; 0 other;
SQ
Query Match 34.8%; Score 2144.4; DB 24; Length 6608;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 4215; Conservative 0; Mismatches 1796; Indels 543; Gaps 45;
QY 3 GCGCGGGGAGTAGAAGAGCCACTGAGCCATGGCGCAGCGAGTCCACCGAGCAGCGG 62
DB 162 GGCTAGCGGGGAGAGTAGTCGGAGGAGCCATGGCGCGGAGCTCCACCGAGCAGCGC 221
QY 63 AGCCCGAGCAGCGCG---GGCGGGAGCGACACGCGCGAGCGAGTGGTCTCAGTGGCCAT 119
DB 222 AGCCCGAGCAGCGCGCGGAGGAGTCCACGCGCGTGTGAGCCGAGCCAGCGGCGGC 281
QY 120 GGGCGCCGAGTGAAGCTCGGGA-----GCAGCTGGAGACCCCGCGAGCGCGGAC 170
DB 282 GGGCCCTCGCGGAGGGGGGCGGACAGACACCGCGGACCCCGCATCGCTGCTCGGAC 341
QY 171 CCGGCCACCAAGCTCCACAGAGAATGGCCAGCTGTCTTCTCAACGGCGTAGCTGAA 230
DB 342 CCGGCCACCAAGCTCTTACAGAGAATGGTCAAGTGTCTTCAACCATCAATGGCGTAGCTGAG 401
QY 231 CAAGGAGATGTCCATGTCCAAAGAGGAAACACGAGGAGG----- 268
DB 402 CAAGATGAGTCAAGCTCCAGAGGGGTGACCTTAATGGCCAGAGAAGGAGCGCTCAACGGT 461
QY 269 -----GCAGGAGGAAGAGTCTGTGATGAGGATGTTGGACAGCGA 308

DB 462 CAAGGAGCCCTAAACAGCCAGGAGGAAGAAAGTCAATGTACGAGGTTGGACAGAGA 521
QY 309 GAGTCAGAAAGATGTGAGAGAAAAGACCGAGTTTGAAGAAATGGCGGCAACTCCACAGCT 368
DB 522 GACTCTGAAGATGTGAGCGAAGAGACTCCGATAAAGAGATGGCTACTAAGTCAGCGGTT 581
QY 369 GTTGAAGATATCAAAAGGATGGCAGGAGGAGACATCAGAAATAATTGAACAGATCCCT 428
DB 582 GTTCACGACATCACAGATGATGGCAGGAGGAGA---ACCGAAATATCGAACAGATTCTCT 638
QY 429 GCTTCAGAAAACAATGTGAAGAAATGGTACAGCCTGCTGAGTCCCGAGGCTAATGATGTT 488
DB 639 TCTTCAGAAAGCAATTTAGAAGAGCTAACACACCCACTGAGTCCCGAGGCTAATGATATT 698
QY 489 GGCTTCAAGAAAGTATTTAAATTTGTGTTTAAATTTACCGTGAAGAGGATAAAAAAT 548
DB 699 GGATTTAAGNAGGTGTTTAAAGTTTGTGCTTTAAATTTCACTGTGAAGAGGATAGACA 758
QY 549 GAAAAGTCAGATACTGTCCAATCTACTCTCAAGAGGATGAAGCGGAGGGGAGAGAA 608
DB 759 GAGAAGCCTGACACTGTCCAGCTACTCTACTGTGAAGAAAGATGAAGGGAGGAGCAGCA 818
QY 609 GCCTCTGTGGAGCTGGAGACCAAGGAGGAGCCAGTGTGGAGACTGCCGTCCGAGAGTCA 668
DB 819 -----GGGGCTGGCGACCAAGGAGCCCGAGGCTTTGGGGCT-----GGAGAGCA 863
QY 669 GCATCCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAAGCAAGAGGACCCCTGAAGCAA 728
DB 864 GCATCCAAAGAAAGCGAACCCAAACCAATCTACAGAGAAACCCGAGAGACCCCTGAAGCGT 923
QY 729 GAACAGAGCAGCAGCAAAATCCCTTCAAGCGCAATCTGATCAAGCGGTGAGGAGAA 788
DB 924 GAGCAAGCCAGCAGCAAAATTTCTCCCGAGCGCAATCTGGCCAAAGCAGT---GGAGAA 980
QY 789 GCCAAGATGAAGGAGAGAAACCAAGAGAAAGAGCCCAAGTCCCGCAGATCCCGG 848
DB 981 TGCAAGAGAGGAGGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTCCG 1040
QY 849 AGCAGCGCCAGTCAACAGTGAAGCAACATCTCTCTCAAGAAAGTCTTCACTCAAGTGG 908
DB 1041 ACTAGTCCCGTGAAGTGAAGCAAGATCACTTCAAAAATTTCTCACTCAAGTGG 1100
QY 909 GCGGCTGGCGCAAGAGACCGAGTTCAGAAATCAAAAAGAGATGATCTGGAATCTGCC 968
DB 1101 GCGGCTGGCGCAAAAAGAACCAAGTTCAGGAAGCCGAGGAGTGAAGTGAAGTTCA 1160
QY 969 GAGAAGAAAGGAGCAAGGAGGAGAAAGTAGACAGGAGAAAGAAAGAAAGAAAGAGAG 1028
DB 1161 GAGAAGAAAGGAGCAAGGAGGAGAAAGTAGACAGGAGAAAGAAAGAAAGAGAGAG 1220
QY 1029 CCAGCTTCGGAGG-----GCAGGAGCGGCGAGAGAC 1061
DB 1221 GTTGCCTTCGAGNAATGACCGCTCCGAGCAAGCCACCCAGAGGCGCGCAGAAAGT 1280
QY 1062 ACAGACCGAGCGCAGGTTGTGTCAGCAGACTACGAGAGGTGGAGCTGCTTTTGAAGACCAG 1121
DB 1281 GCCCAGAGCGCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGAGCA 1340
QY 1122 GTTGTGACCTTGAGGAGCATGTCAGAGGAGAGTGTGCTCTTTTGGCAACGGAAGTGT 1181
DB 1341 GTCAGTGGCTCGCAGGAGACCTTCTGAAGAGAAACCTGCTCCGTTGGCGCAGAGAGT 1400
QY 1182 GATGAGAAGATGAAGCCCAAGAA---GTTGTTGACAGAGTCCACCTGAGCAGCCGTG 1238
DB 1401 GATGAGAAGATGAAGTCCCAAGAAAGAGTGTGGCGAAGTCCACCTGAGCAGCCGTG 1460
QY 1239 GAGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1298
DB 1461 GAGGAGAGAACCGAAGAGCAGAAA-----ACGGAGGTGGAA 1496
QY 1299 GGAACAGGAGAGATCTTGGCCCCCTGAGAAATCTGGCTGAGCCCCCAGAGAGTCCCCCAGGAA 1358

Db 1497 GAAACAGAGGCTGTGCCAGCTGAGAAATTGGTGAATGATGCAAGAACTCAGSAA 1556
 Qy 1359 GCTGAGCCTGTGAGAGCTGATGAAGAGCAGAGATGTGTCTCTGAGAGAACAC 1418
 Db 1557 GCCGACCTGCGAAGAGAGCTGTGAAGCTCAAGAAAGTGTGTTCCGAGAGAGACCTT 1616
 Qy 1419 ACTCACTGACAGACCTTAAGTCTTGAGAGAGAACGCTGCCAAACACCGAAGGACTT 1478
 Db 1617 ACACAGGAGAGCTGACCTCAGTCTGATGAGAAAGTGTCTCAAAACCCCGAAGGCTT 1676
 Qy 1479 GTCAGTAGGAGGAGATGCTGCTCTCAGGAAAGATCAAGGTACAGAGAGTCTCTTG 1538
 Db 1677 GTGAGTAGGAGGAGATGCTGTCTCATCAGAGAGAAATGAAGTGCAGAGAGTCACTA 1736
 Qy 1539 AAGAACTCTTCAGTAGCTCAGGCTTAAAGAACTGTCTGGAGAGAGCAGAGGGGAAA 1598
 Db 1737 AAGAAAGCTTTTACAGACAGCTGGCTTAAAGAAAGCTTCTGGAAAGAAACAGAAAGGAAA 1796
 Qy 1599 CGAGAGGAGTGGGAGAGAGAGAGCTTGAGAAATCAACACATTTCAACCGAATCCCA 1658
 Db 1797 AGAGGA---GGAGAGAGAGAGAAATCAGGGGAGCAACTCAGGTTCCAGCCGATTTCTCG 1853
 Qy 1659 GAGAGTGTGATGAGACAGAAAGGAGAGAGCTGTGCTGCTCCCGAGAGAGCTGAGAGAG 1718
 Db 1854 GACAGCCAGAGAGAGAGAGAGAGAGAGAGCTGTGCTCTCATCTCCCTGAGAGAGCCGAGAGAG 1913
 Qy 1719 ACCAGTGTCTGAGAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1778
 Db 1914 ATCAGTGTCTGAGAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1973
 Qy 1779 ACTACTCTCGATGAG 1838
 Db 1974 GCTACTCTCGATGAG 2033
 Qy 1839 AAGTGAG 1898
 Db 2034 ATGAG 2093
 Qy 1899 GAGAGAGTGAAG 1958
 Db 2094 GACAGAGTGAAG 2153
 Qy 1959 GAAGTCAAAACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2018
 Db 2154 GAATGAAAG 2213
 Qy 2019 TCAGTCTCTTGGAG 2078
 Db 2214 TCAGTATCTTGGAG 2273
 Qy 2079 TCTCTTTCAGATGATGAG 2138
 Db 2274 TCTCTTTCATGATGAG 2233
 Qy 2139 GAGGAG 2198
 Db 2334 GAGGAG 2293
 Qy 2199 GAGGAG 2258
 Db 2394 GATTCAG 2453
 Qy 2259 GTCTCCAGTGGAG 2318
 Db 2454 GTTTCACCTGGAG 2513
 Qy 2319 CTGAG 2366
 Db 2514 CTGAG 2573
 Qy 2367 ATGAG 2426
 Db 2574 ACTGAACTCCGTTAAAG 2633

Qy 2427 AAGAAAGGAG 2486
 Db 2634 AAGAAAGGAG 2693
 Qy 2487 ATAAATGAG 2546
 Db 2694 GCCAAG 2753
 Qy 2547 GAG 2597
 Db 2754 GAAAG 2813
 Qy 2598 GGGGCTGTGATGAG 2657
 Db 2814 GCAGCAG 2873
 Qy 2658 GTCAATTGATGAG 2717
 Db 2874 GTGCTGAG 2933
 Qy 2718 GCTTCCTGATCAG 2777
 Db 2934 GCTTCAGTGAAG 2993
 Qy 2778 GTCACGTAAG 2834
 Db 2994 GTATTTGAG 3053
 Qy 2835 GAGGATTAAG 2894
 Db 3054 GAGAAAG 3113
 Qy 2895 GTGACAG 2954
 Db 3114 GTGACAGTGAAG 3173
 Qy 2955 GCCGAG 3014
 Db 3174 GCTGAAG 3233
 Qy 3015 ACAG 3074
 Db 3234 ACAG 3293
 Qy 3075 GAGGAG 3134
 Db 3294 GAGAGGAG 3353
 Qy 3135 GTGCTGAG 3191
 Db 3354 GTGCTGAG 3413
 Qy 3192 GAGGTAG 3251
 Db 3414 AAG 3473
 Qy 3252 GAGAGGAG 3311
 Db 3474 GTAGATGCTCAG 3533
 Qy 3312 ACTCAG 3359
 Db 3534 ACCCAG 3593
 Qy 3360 GTGAGCAG 3401
 Db 3594 GTTACAG 3653
 Qy 3402 CAGGCGGTGAG 3461
 Db 3654 CAGGCTATCCCTGAG 3713

CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 6608 BP; 2091 A; 1418 C; 1827 G; 1272 T; 0 other;

Query Match 34.8%; Score 2144.4; DB 24; Length 6608;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 4215; Conservative 0; Mismatches 1796; Indels 543; Gaps 45;

QY	3	GGCGGGGAGTAGAGGCCACTGAGCCATGGCGCCAGCGAGTCCACCGGACGCGG	62
DB	162	GGTAGCGGGAGAAAGTGGGAGGAGCCATGGCGCCGAGAGTCCACCGAGCAGCGC	221
QY	63	AGCCCGGAGCAGCGG-----GGGGGAGGACACGCGGAGCGAGTGGTCTCAGTGGCCAT	119
DB	222	AGCCCGGAGCAGCGCGCGGAGGGAGTCCACGCGGCTGAGCCCGAGCCAGCGGCGC	281
QY	120	GGGCGCGCAGCTGAAGCCTCGGA-----GCAGCTGGAGACCCCGCGAGCGGAC	170
DB	282	GGGCGCCTCGCGGAGGGCGGCGCAGACACCGCGGACCCCGCATCGCTCGCTCGGAC	341
QY	171	CCGCGCACCAAGCTCCACAGAGAAATGGCCAGCTGTCTCTCAACGGGAGTGTAA	230
DB	342	CCGCGCACCAAGCTCTACAGAGAAATGGTCAGCTGTCCACCATCAATGGCGTAGT	401
QY	231	CAAGGAGATGTCCATGTCCCAAGAGGAAACACGAGGG-----	268
DB	402	CAAGATGAGCTCAGCTCCAGGAGGGTGACCTAAATGGCCAGAAAGGAGCGCTGA	461
QY	269	-----GCAGGAGGAAGAGTCTGTTGATGAGGATGTTGGACACGCA	308
DB	462	CAAGGAGCCCTAAACAGCCAGGAGGAAGAGTCTTGTACGGAGGTTGGACAGAG	521
QY	309	GAGTCAGAGATGTGAGAGAAAGACCGAGTTGAAGAAATGGCGGCAACTCCAGCT	368
DB	522	GACTCTGAAGATGTGACGAGAAAGAGACTCCGATAAAGAGATGGCTACTAAAT	581
QY	369	GTTGAAGATATCAAAAGGATGGCAGGAGAGACATCAGAAATTAATTGAACAGAT	428
DB	582	GTTACGACATCAGATGATGGCAGGAGGAG-----ACCGAATATCGAACAGAT	638
QY	429	GCTTCAGAAAACATGTGGAAGAAATGGTACAGCTGCTCAGTCCAGGCTTAATG	488
DB	639	TCTTCAGAAAGCAATTTAGAAGAGCTAAACACACCCACTCAGTCCCAAGGCTA	698
QY	489	GGCTTCAAGAAATTTAAATTTGTTGGTTTAAATTCAGGTTGAAGAGGATAAA	548
DB	699	GGATTTAAGAAAGTGTAAAGTTGTTGGCTTTAAATTTCACTGTGAAAAAGGA	758
QY	549	GAAGAAAGTCAGATCTGTCACCTACTCTCAAGAGGATGAAGCGGAGGGGAGAA	608
DB	759	GAGAAAGCTGACACTGTCCAGCTACTCCTGTGAGAAAGATGAAGGGAGGAGCA	818
QY	609	GCCTTGTCCGAGTGGAGACCAACAGGAGCCAGTGTGGAGACTGCGGTCCGAG	668
DB	819	-----GGGGGTGGCGCACCCAGGACCCAGCCTTGGGCT-----GGAGA	863
QY	669	GCATCCAAAGAAAGTGAAGCAATCCACAGAGAGCAAGAGGACCCCTGAAGCAA	728
DB	864	GCATCCAAAGAAAGCGCAACCCAAACAAATCTACAGAGAAACCCGAGAGAC	923
QY	729	GAACAGAGCAGCAGAAATCCCGCTTCAAGCGCAATCTGATCAAGGGGCTGA	788
DB	924	GAGCAAGCCACGAGAAATTTCTCCCGCAGCGCAATCTGGCCAGAGCT-----	980
QY	789	GCCAAAGATGAAGGAGAGAAACCAAGAGAAAGAGCCCAAGTCCCGCAGAAAT	848
DB	981	TGCAAGAGGAGGAGAGAGAAACCAAGAAAGAGAAAGAGAAAGAGAGAGAG	1040
QY	849	AGCAGCCAGTCAACAGTAGAGCAACATCTCTCTTCAAGAAAGTTCTTCACT	908

DB	1041	ACTAGTCCCGTGACCGATGAAACAGGATCAACCTTCAAAAAATCTTCACTCAAG	1100
QY	909	GC CGGTGGCGCAAGAACCAAGCTTTCAAGAAATCAAAAGAGGATGATCTGGA	968
DB	1101	GC CGGTGGCGCAAAAGAACCAAGCTTTCAAGAACCCGAGGAGGATGAGTGA	1160
QY	969	GAGAGAGAAAGGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1028
DB	1161	GAGAGAGAAAGGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1220
QY	1029	CCAGCCTCGGAGG-----GCAGGAGCGGAGGAGGAGGAGGAGGAGGAGG	1061
DB	1221	GTTGCTTCGAGAAACTGACCGCTCCGAGCAAGCCACCCACAGGAGCGGAG	1280
QY	1062	ACAGACCGAGCCAGGTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1121
DB	1281	GCCACGAGCGCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAG	1340
QY	1122	GTTGTTGACCTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1181
DB	1341	GTCAGTGGCTCGCAGGAGCCTTCTGAAGAGAAACCTGCTCGTTGGCGACA	1400
QY	1182	GATGAGAGATGGAAGCCACCAAGAA-----GTTGTTGAGAGGCTCCAGT	1238
DB	1401	GATGAGAAATAGAGTCCACCAAGAGAGGTTGTGGCGAAGTCCACGTCAG	1460
QY	1239	GAGAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1298
DB	1461	GAGGAGAAACGAGAGCAGAA-----ACGAGAGTGGAA	1496
QY	1299	GGAACAGGAGAAATCTTTGCCCCCTGAGAACTGGCTGAGCCCCAGGAGGT	1358
DB	1497	GAAACAGCAGGGTCTGTGCGAGTGAAGAAATGTTGGAATGATGAGAACT	1556
QY	1359	GCTGAGCTGCTGAGGAGTGTGAGAGCAGAGAGATGTGTCTCTGAGAGGAC	1418
DB	1557	GCCGAACCTGCCAAGGAGCTGGAAGCTCAAAGAAACGTTGTTCCGAGAGG	1616
QY	1419	ACTCAACTGACAGACCTTAAGTCTGGAAGAGAGACCTGCCAAACACCC	1478
DB	1617	ACACAGGAGTGAACCTCAGTCTGATGAGAGGCTGTCTCAAAACCCCGG	1676
QY	1479	GTCAAGTGGAGTGTGCTCTCTCAGGAAAGAAATCAAGGTACAGGGAAGT	1538
DB	1677	GTGAGTGGAGTGAATGCTGTCTCATCACAGGAGAGAAATGAAGCTGCA	1736
QY	1539	AAGAACTCTTCAAGTGTGAGCTTAAAGAGCTGTCTGGGAGAGGAGGAGG	1598
DB	1737	AAGAACTTTTACCAGCCTGCTTTAAAGAAAGCTTTCTGGAAGAAACAG	1796
QY	1599	CGAGGAGGTGGGGAGACGAAGAGCTGAGAAATACCAACATTCACACCG	1658
DB	1797	AGAGG-----GGAGAGACGAGGAAATCAGGGAGCACACTCAGGTTCA	1853
QY	1659	GAGAGTGTCTGATGAGCAGAAAGGAGAGAGCTCTGCGTCTGCTCCCGG	1718
DB	1854	GACAGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCTCTCATCTCCCT	1913
QY	1719	ACCACGTGTCTGAGAGAAAGGGCGCTGGAAGCAACCCAGGATGGGGA	1778
DB	1914	ATCACGTGTCTGGAAGAGGGCTTAGCCGAGGTGACAGAGTGGGGAAGG	1973
QY	1779	ACTACTTCGATGGAGAGAAAGAGAGAGGATCACTCTCTGGGCATCTT	1838
DB	1974	GCTACTTCGATGGAGAGAAAGAGAGAGGTTGTCCTCTCTGGGCATCTT	2033
QY	1839	ATGGTGACACCCAGAAACGGGTCCGAGACCTTCTGAGAGTCAAGGAGG	1898
DB	2034	ATGGTGACCCCAAGAGAGCGTGTAGACGCGCTTCCGAAAGTGAATAA	2093
QY	1899	GAGAGGTCAAGAGCGCCACCTTGTCTCTCACTGATAGCAGTGTCAAGAA	1958

Db 2094 GACAAAGTCAAAGCGCTACCTTGTCTTCCACCGAGACACAGCTCTGAATGCAAGAA 2153
Qy 1959 GAAGTCAAACTGTTGGTAGAGAACTAAAGCCAGAGAAACCAAGCGTAGGTGATATCT 2018
Db 2154 GAAATGAAGGAGACGTGGAGAGCGAAGCCGGAAGAACCAAGCGCAAGTGGATCC 2213
Qy 2019 TCAGTGTCTTGGGAGCACTGATTTGTGTGGATCATCAAGAGAGCAAGAGCA 2078
Db 2214 TCAGTATCTTGGGAGCTTATTTGTGTGGATCATCAAGAGAGCAAGAGCAAG 2273
Qy 2079 TCCCTTTCAGATGATGAGAGGCGCAAGCACTGGAGGGGACATGCACAGACAG 2138
Db 2274 TCCCTTTCAGATGAGAGGCGCAAGCACTGGAGGGGACATGCACAGACAG 2333
Qy 2139 GAGGCGAGCAAGCAAGAGAGCCGGAACAGACGCTGTTCTGCGACGACCCAGAGCAG 2198
Db 2334 GAGGCGAGCAAGCAAGAGAGCCGGAACAGACGCGGATCTTGTGCTGTTCCAAAGCAT 2393
Qy 2199 GACCAAGCGCAAGAGATTTCTCACCCGAGCCAGCGGAGAGCCCTTCCGAAGGGAGGT 2258
Db 2394 GATCCAGGGCAGGGAGTTCCTCCCGGAGCAAGCTGGAGCCCTTACCGAAGGGAGGGC 2453
Qy 2259 GTCTCCACTTGGGAGTCAATTAAGATTAGTCACTCCAGAAAAATCCAAATCAAAA 2318
Db 2454 GTTTCACCTGGAGTCAATTAAGATTAGTCACTCCAGAAAAATCCAAATCCAAAG 2513
Qy 2319 CTGGAAGAGAAAGCCGGAAGC-----TCTAGTGAAGCAAGTGTCTCACTGAG 2366
Db 2514 CTGGAAGAGAAAGCCGGAAGCCTCCATAGCTGGGTCTGGGTGTAAGCAATTCACCTCAGAC 2573
Qy 2367 ATCGAACCGATGAGAAAGATCTTGGGTTTCCATTAGAAATTCATCCCGGAGGGCG 2426
Db 2574 ACTGAACCGGATGAGAAAGATCTGGGTCTCAATCAAGAGTTATTCCTGAGAGAAAG 2633
Qy 2427 AAGAAAAGGCGAGAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2486
Db 2634 AAGAAAAGGCGAGAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2693
Qy 2487 AATAATGAGAGCAAGCCCTATGTCCTGAGCCCTGTCCTGTCGATGATATATGCACTG 2546
Db 2694 GCCAAGAGATGATCTGTATGTCCCGGCGTGTGTCCTGTCGATGATATGTCGTA 2753
Qy 2547 GAGAGGAGAGATG-----AAGCCAGGGGAAATACGAGAGTGCCTCCAGCTGCTG 2597
Db 2754 GAAAGGAGAGAAATGAGGAGCAAGCAAGCCCAAAAGGCGCAGAGAGCCGAGCAAG 2813
Qy 2598 GGGGCTGTGATGTCCTCCAGAGAGCTCAATGAGCTGTGTCCACACTGATGAGTCGA 2657
Db 2814 GAGGCACTGAGGTGTCCAAAGAGCTCAGAGAGTCAAGGTTCAATGATGCGACAGCT 2873
Qy 2658 GTCATGATGAGGAGCCAGGAGCTCAGAGTGTGAAAGAGCGGTCTCTTGTGGATATCC 2717
Db 2874 GTCCGTCGAGGAGCGAGGAGCTCAATTAATGAAAGAGGTCTCTTGTGGATATCT 2933
Qy 2718 GCTTCCGTAACAGAACTCTTGAACAACAAGCGGAGAAAGCCATGCACTGTTGAAG 2777
Db 2934 GCTTCAGTACAGAACTCTTGAACAAGTGAAGCTGAAGCCCACTGTTAACTGAGAG 2993
Qy 2778 GTCACTGAAAAAGACATCAATGTC---AGAAGAACTCTGTGCTCAACCAAGCTTACCA 2834
Db 2994 GTATTGAAAAAGAAATGATTAATGCAAGAAAGAAACCCCAAGGTTACTGAACTCTGCCA 3053
Qy 2835 GAGGTTAAAGATGCCATGACGACATGTCACCAAGTGAAGTGAATTTCACTGAGAGCT 2894
Db 3054 GAGAAACAGAGAGCCCGGGGCGACAGGTCTGTAAGTGAAGCGGAATTTGACCCCGAAGCT 3113
Qy 2895 GTGACAGCCACAGAGACTTCAGAGGCTCTCCGTACTGAAGAGTTACCAAGCATGGGG 2954
Db 3114 GTGACAGCTGCAAGAACTGACAGGCGCATTTGGTTCCGAAGAAAGAAACCGAAGCATCTGCT 3173
Qy 2955 GCCGAAGAGACCAAGACATGTTGTCCGAGTTTCCAGTGACTGATCCCGAGACCC 3014
Db 3174 GCTGAAGAGACCAAGAAATGTTGTCAAGAGTCTCCAGTTAAACCGACTCCCGAGACCC 3233

Qy 3015 ACAGAGAGACCAACCCAGTTCAAGAGGTAGAGAGTGTGTCTAGATACAGAGAGAG 3074
Db 3234 ACAGAGAGACCACTCCGGTGCAGAGAGTGAAGGTGGCTTACCTGACATGAAGAGCA 3293
Qy 3075 GAGGCGAGACGAGAGCCATCTCTCAAGCCCTTGCAGACAGGTGAAGAGAGTCCAG 3134
Db 3294 GAGAGGGGAGCTAAGAGGTCTCTCAGGCGAGTGGCAAGAAAGTGAAGAGGATCCAG 3353
Qy 3135 GTGCTTCAACCC---CAGACTGTGCAGAGAACGGGCTCAAAAGCATCTGGAGAGTTGAG 3191
Db 3354 CTGCTGGCACCGGTGGCGCAAGAGATGTGCTTCAAGCTGTGCAGAGAGAGCGCAGAA 3413
Qy 3192 GAGGTAGAGAGAGCTCCGAAGTGTGGCTTCCGAGAAAGAGAGAGCTTATGCCGAA 3251
Db 3414 AGACCAAGAGAGAGCTTGAAGGCTCGGGTCTGAAGAAAGAGAGATGATGTTGAAA 3473
Qy 3252 GAGCCCGTGCAGAGAGCTGAGAGTGAACATCTTGCACAGGGCTCTGAGACTGACAGGCT 3311
Db 3474 GTAGATGCTCAGAGAGGCAAAAACCTGAGCTTTTACAGAGGAAAGTGTGGGCAAGCC 3533
Qy 3312 ACTCAGAGAGCTTGG---AAGTTCTGAAGTCAAGCAATGTGAGC-----CAT 3359
Db 3534 ACCCAGAAAGCTTTGAAAAAGCTCTCAAGTCAAGAGACATAGAGTCCAGTGAAGCTT 3593
Qy 3360 GTCCGCAAGTGGCA-----GGTTATCAAGCTCCAGAGCTGATGAAA 3401
Db 3594 GTAACCACTGTCAAGCCGAAACCTTAGCTGGGTAAATATCAAGAGATGTATGAAA 3653
Qy 3402 CAGGCGGTGGCCCTGAGTCAATCCGAAACCTTGAAGAGAGTGAAGAGCACT 3461
Db 3654 CAGGCTATCCCGCTGACTGGTGAAGAACCCCTTACAGAGTGAAGTGAAGAGCAC 3713
Qy 3462 CCTTTAGCAATTCAGACATGCAAGATGGACACAGCAAGTGAACCATTTGACAGCAG 3521
Db 3714 CCGTAGCCGACTTTGACGACACAGAGCAAAACCAAGAAAGCAGATTTGTGAAATCAT 3773
Qy 3522 GACAGTAAGCACTGCAAGCTGAGAGTCAAGAGTCAAGAGTAAGAGAGAGCGGCTACT 3581
Db 3774 GAGGAGATGAGTGTGCAATCTGTGTACCAAGGAGGCAAGAGAGAGAGAGAGTCTT 3833
Qy 3582 GCTCAGAAAGAGAGCCCTTGCACACTAATTAATGTTCCAGCCAGAGAAACATGGG 3641
Db 3834 GCACGAAAGAGAGGCTCTCAGAGCTTCCAGTTTGTGTCCAGAGAAAGAACTTAAGAA 3893
Qy 3642 GAAAGAACAGAGAGATGTTCTTGAACCTTACACAGCAAGCTTACTGTCAGCCGTG 3701
Db 3894 CAATCAAGATGAGAAACACTCTAGAGCATACAGTTAAAGAGTGTCAGTGAAAATGTA 3953
Qy 3702 CCGCTTTCGCAANAAGTGAAGTGTGCAAGAGGTGA-----GTTGACTGTTG 3752
Db 3954 TCCATTTGTCAAAAGCTGAGGGGACTCAAGAGCTGACAGATGCTGATGAGAAAAC 4013
Qy 3753 GATGAGAAAAAAGTCAAGAAAGACAGAGGTGTTGTACACTGTG-----ACCCAAC 3806
Db 4014 AAGAGACATCACTTTTGAAGAGACTTGAAGGGCTTATAGACAGAGCATTAACATGAGT 4073
Qy 3807 AGTCAAAAGGCTGTGATGTGACATATGACATGAAGTGAAGTGAAGTGGCCGGGTGTAG 3866
Db 4074 CCGGAAAAAGTCACTGAAGTGGCCCTTAAAGGTGAAGGAGACAGAAAGCTGAATGTAAA 4133
Qy 3867 GAAAAAGAGATACAGAGTGCAGAG-----TCTTAAGCTGAGAGAGGA 3911
Db 4134 AAGATGATGCTCTTGAACCTGAGAGTCAAGCTTAAGTCTCTCTCAATCCCGTGAAGAG 4193
Qy 3912 GAGATGAAAACTGACGTTGAAGAGAGAAAGAGAGACAAAGCCAGAGCAAGTGAAGTA 3971
Db 4194 GAGATGTGATGTTCAAGTGAAGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 4253
Qy 3972 GAAG---TGAGCAGAGAAACAGCCGCTCTGAGCATGAAGAGAACTTACGGAGAGCAGTTC 4028
Db 4254 GAGAGGCTTGAACAGAAACAGCTGTATCTGAAGAGAGTCAAGTAAAGAGCTCTCTC 4313

QY 4029 CTGACACTTGACATGCCAGCTCAGAGGGGGAGGCACTGGGAAGCCTTGGAGGAGC 4088
DB 4314 CAGACAGTGAATGTGCCCATCATAGATGGGGCAAGGAAGTCAGCAGCTTTGGAGGAGC 4373
QY 4089 CCTTCTC---TCCAGACCAAGACAAGAGCAGGTTCATAGAGTTCAAGTTCAAGCCTG 4145
DB 4374 CCTCCTCCCTGCTAGCTCAAGAGGAGCAGTAGTGCACCAAAATTCAGTTTCAGAGTCT 4433
QY 4146 GACACAACTGCTCACTAAAACAGCAGAAAGCTGTGGAAAA-----GGTCATAGAAACGGTT 4199
DB 4434 GAGGCATCATTCCTCTAAACAGCGGCTGCAGAGGAGGAAAAAGGCTCTTAGGAGAACTGCC 4493
QY 4200 GTGATTTTCAGACAGAGTGAAGTCCAGAGTGTCTAGTGCACACTTTATTACAGCTGAG 4259
DB 4494 AACATTTTGAAGAACAGGTGAACCTTTGGAGCCTGCAGGTGCACATTTAGTTCTGGAAAG 4553
QY 4260 AAGTCTCTGCAACGGGTGGCCACTGACCTCTTCAGCATGCAGAGGACAGGTACCCCTG 4319
DB 4554 AATCCTCTGAAAAAATGAAGACTTTTGGCGCTCATCCAGGGGAGATGCTGTGCCACA 4613
QY 4320 GGGCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAATCTCTGCTCTCTGAAAGCACC 4379
DB 4614 GGGCCGACTGTGAGGCAAAATCGACACCAAGTATGATCTCTGCTACTACCAAGAAAGGC 4673
QY 4380 CTACATCTGACCTACAGCAGAAATGAAGCGCATCCAGAGAGCGCATCAGAGGAAGAG 4439
DB 4674 TTAAGTTCCGACCTGGAGAGAGAAACCAATCACTGAAGTGGAAAGTCAGATGAATC 4733
QY 4440 GACAAGCCAGATGCTGCTGCTGATGCTGACGGCAAGGAGAGTACAGCAATCGAAAAAGTC 4499
DB 4734 GATGAGCAGTTGCTGCC-----AGGAGTCAAAGTGAGTGAAGTGTG---AGGAT 4784
QY 4500 CTCAAGCTGAACCTGAGATCTCTGGAACTTGAGAGTGAAGCAACAAGATTTGCTGAAC 4559
DB 4785 TTAGAGCCTGAAAATGGATTTTGGAACTTGAGACCAAAAGCAGTAAACTTGTCCAAAAC 4844
QY 4560 GTCATTCAGACAGCCGTTGACCAAGTTCGACAGTAC-----AGAAACAGCCCCGAAACTCAT 4616
DB 4845 ATCATCAGACAGCCGTTGACCAAGTTGTACGTACAGAGAAACAGCCACCGAAATGTTG 4904
QY 4617 GCTTATGATTACAGACCCAGGTTCT-----CTGCATGCAGGCTTGACAG 4660
DB 4905 ACCTCTGAGTTACAGACACAAGCTCAGTGATAAAGCTCACAGCCAGGAGCGCTGGACAG 4964
QY 4661 CAGGGAGCCCAACAGATGCTGGCAAAATGAAAGATGCCAAGATGAAACACCCAGTGCC 4720
DB 4965 GAAACGAGAAAGAGAGAGGAACTCAGGCTCTGCACAGGATGAAACACCAATTAAT 5024
QY 4721 GCAGCCCA-----GAGAGGACTTGCAAGTCTGACCGTTCTGGA 4759
DB 5025 TCAGCCAAAGAGGAGTCAAGAGTCAACCCGAGTGGGACAAGCAATCTCGATATTTCCAAA 5084
QY 4760 GGCATGGCTCAGCCTCGGAAATGCTTGCCGCGCTTGCACTTGAAAGCGCGGTGTCAA- 4818
DB 5085 GACATGAGTGAAGCCTCAGAAAAAGACCATGACTGTTGAGGTAGAAGTTCCACTGTAAAT 5144
QY 4819 ----- 4818
DB 5145 GATCAGCAGCTGGAAGAGTTCGTCCTCCCATCTGAGGAAGAGGAGGTGGAGCTGGNACA 5204
QY 4819 -----AGTAAGCATGTGAGAGCTGCCT 4840
DB 5205 AAGTCTGTGCAGAAAGATGATGTGCTCATGCTTGTGTAGCAGAAAGAAATAGAGAACTCACTA 5264
QY 4841 CCTCAACCCCAAGA--TCCAAAAGGAGCATGCTGCTGATGGCCCTCAGCTCCAAAGCTTA 4898
DB 5265 GTTGAACCGAAGAAAGATGAAAAGGTGATGTGATCACCCTGAAAAACCCAGAACTCA 5324
QY 4899 GCCCAGCAGAGGCGAGTCTCTGGAAAACCTAACCAAAAGAAATCCCCAGACACCAACGGA 4958
DB 5325 GCCCTGGCTGATACTGATGCTCAGGAGGCTTAACCAAGAGTCCCCAGATACAAATGGA 5384
QY 4959 CCAAAGCTAACCGAGGAGGCGGATCCCCCAAAAGTTGAGGTCCAGGAAGAAGAAATGTCT 5018

DB 5385 CCAAAACAAAAGAGAGAGGAGGATGCCAGGAAGTAGAATTGCAGGAAGAAAAGTGCAC 5444
QY 5019 ACCAAGTCAAGTCAAAGAGAACAAAGGCCAGGCAAGAGAGACCTGCAGGAGCCAAAGGGA 5078
DB 5445 AGTGAATCAGATAAAGGATCACACCCCAAGACACAGAGGAGTTACAGAAACAAGAGAGA 5504
QY 5079 GACCTGGCAGAACTCCPAAGATGTTAGTTG-----CTCATTTGTAC 5117
DB 5505 GAATCTGCAAGTCAGAACTTACAGAACTTAAACCATCATGCAGTTAAACTCATTTGTCT 5564
QY 5118 ATCTGTAAGACCAAGATGTGAAACAAGTACAGAACAA-----GATGCTGCTTTGGGAC 5173
DB 5565 GTTTGGAAGACCAAGATGTGAAGACAAGTAGTAGAAGAAAATGAATGCTGCTGCTGAGA- 5623
QY 5174 CTTGAGACCAAGATTTTCAGAGCCCATGAGATCCAGAGCAGGCGCTCCAAATGATTTCC 5233
DB 5624 CTGAAGCCAGTATTTTCAG-ACTTTGAGNAATTGGAGAGCAGGCACATCACTGATCTCA 5682
QY 5234 ACCAGTAGACACCCCGACAATTCAGAGCTTCATCGGAGCTAGAGCCAGCTAACATT 5293
DB 5683 TTTCTAGAGAGC-CCCTGACAATCCTGAGGCTTCATCAGAGCTAGAGCCATTTAACATT 5741
QY 5294 TCCTCGTTTCAAGACTGCCTTTGATTTGGCCCTTGATGCGTCCGTGATTTCTAACATT 5353
DB 5742 TCCTCTTTCCAGACCAACCTACAATTTTCCCTTGATAAC----- 5781
QY 5354 TCCTCGTTTCAAGACTGCCTTTGATTTGGCCCTTGATGCGTCCGTGATTTCCGGAATTA 5413
DB 5782 -----CATATAAATTCGATTTA 5799
QY 5414 AGTCTCGCTTCTCAACCTGGAAAC-CAATTTGCCATACCTAGTTCCCACTTCTCAAACT 5472
DB 5800 AGTCTCTAAATCTTAACCTGGAACTGGAGTTGGCAATACCTAGTTCTGCTTCTGAACT 5859
QY 5473 GGAGCATCTCCTTTATGATTTATATGATGTTTATGATGCTTCTGCTGCTGCTGCTAT 5532
DB 5860 GGAGTATCATCTTTACATATTTATATGATGTTTAAAGTAG--TCCTCCTGATCTAT 5916
QY 5533 TGTATATTTTCTTCAAGCTT-----AAGCACATGCTTTTGTATTA 5575
DB 5917 TGTATATTTTCTTAAATGTTTAAAGAAATGTGAGGAATACATGATGCTTTTGTATCA 5976
QY 5576 TGCAATATATAACGGGTGTGCAGCCATAGCAGCTTTGAAAAGCTCCAGCCTCAACTG 5635
DB 5977 CACAGTATATGATGGGCGCATGTGCCATAGTGCAGGCTTGGGAGCTTTAAGCCTCAGTTA 6036
QY 5636 TAACTGTCAGCAACAGATAAATTC-----CTGGCAAGAGAGACAAGCTTTTTT 5686
DB 6037 TATAACCCCAAAAAACAGAGCCTCCTAGATGTAAACATTCCTGATCAAGGTACAATTTCT 6096
QY 5687 TAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGTAAGAGTGGTTGTTTCTCT 5746
DB 6097 TAAATTTCTACTAATGATGAGGTCCATATTTAGTGTGATCTCTGAAAATGTTGCTTTCTCT 6156
QY 5747 ATGCACAGCGAGCTCAGAAAATAAAAAACCCCATTTTGAACATCCAGATGTCCCAATATT 5806
DB 6157 ATTACAGGAGTGTGCCAAAATTAAGAAAGATTTTGAACATACAGAAATGTTCTATTGTC 6216
QY 5807 ACCATGATTTTTCCTCTTTTGTCTAATCCAGTCCAGTCCAGTTGGAAGAGTCTCCTCTG 5866
DB 6217 ATTGGAAAATTTTGC-----TTTCTAACCCAGTGGAGGTTAGAAAAGATTTATTTCT 6269
QY 5867 TGTTCAG-----TTAAGCCCTGCTCTCTTAATGATATGCACAAATGAGT 5909
DB 6270 GGTAGCAAAATTAACCTTTTTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6329
QY 5910 GTGCTTAAAGGCCATGAGA-----TGTTTCTTAATCAGAAAGAACTCTGT 5953
DB 6330 GTGCTTAACTCCTGAGCAAGTAGTGAATATGTTTATATGTTATGAAGAAAGAAATGTT 6389
QY 5954 TGTACCTTTTGTGATGTTACTCTTCTATGCTGGACCGAAATTCATATGCAGATCGAAGTG 6013

Db 6390 TGTAA--TTTGTGATTCTACTCTTATATGCGAGCTGCATTCAACATGCGATGAATA 6447
 QY 6014 AGTCTGTTCTTTACAGATGATTTTGTATGATCTGAGTTGTCTGTATATCTG 6073
 Db 6448 AGTCAGGTTCTTTCAAAATGATTTTGTATGATCTGAGTTGTGTGTGTCATATTTG 6507
 QY 6074 TGGCCCTCTTTAGAAACATGTCATTATGTCCTTTGATTAATTGTGATTGACAA 6133
 Db 6508 TGCATTCCTTTAGAACATGTTGCAACACATTCATTGATTAAGTTGATTTGACGA 6567
 QY 6134 CTGATTTAAATAAA 6147
 Db 6568 CTGATTTAAATAAA 6581

RESULT 6
 AAV23545
 ID AAV23545 standard; cDNA; 6605 BP.
 AC AAV23545;
 DT 13-JUL-1998 (first entry)
 XX Human gravin polypeptide coding sequence.
 DE
 KW Gravin, kinase anchoring protein; type II regulatory subunit; PKA; PKC;
 KW CAMP-dependent protein kinase; protein kinase C; autoimmune disease;
 KW Myasthenia gravis; nicotinic acetylcholine receptor; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT 192..5534
 FT CDS /*tag= a
 PN US5741890-A.
 PD 21-APR-1998.
 XX 19-DEC-1996; 96US-0769309.
 PF 19-DEC-1996; 96US-0769309.
 XX 19-DEC-1996; 96US-0769309.
 PR (UYOR-) UNIV OREGON HEALTH SCI.
 PA Klaufek TM, Nauert JB, Scott JD;
 PI WPI, 1998-260552/23.
 DR P-PSDB; AAW53863.
 XX New polypeptide fragments of protein kinase binding protein -
 PT are useful for the study of modulation of action between gravin and
 PT protein kinase(s)
 PS Example 1; Column 19-32; 32bp; English.
 XX This sequence encodes the human gravin polypeptide, which represents
 CC a polypeptide of the invention. The polypeptides are fragments capable of
 CC binding to type II regulatory subunit of CAMP-dependent protein kinase
 CC (PKA). Gravin is a kinase anchoring protein that binds to type II
 CC regulatory subunits of PKA and protein kinase C (PKC). Gravin is also an
 CC antigen of the autoimmune disease Myasthenia gravis (MG), where a patient
 CC develops antibodies against their own nicotinic acetylcholine receptors.
 CC The polypeptides are useful for providing analogues of gravin in the
 CC study of the modulation (e.g. blocking, inhibiting and stimulating) of
 CC interactions between gravin and kinase. The peptides are involved in the
 CC modulation of gravin-kinase interactions.
 CC
 XX Sequence 6605 BP; 2091 A; 1416 C; 1826 G; 1272 T; 0 other;
 SQ

Query Match 34.4%; Score 2118.8; DB 19; Length 6605;
 Best Local Similarity 64.2%; Pred. No. 0;
 Matches 4206; Conservative 0; Mismatches 1802; Indels 546; Gaps 46;

QY 3 GGCCGCGGAGTGAAGAGCACTGAGCCATGGGCGCAGAGTTCACCGAGCAGCG 62
 Db 162 GGTAAAGCGCGGGAGAAAGTGGAGAGCCATGGGCGCCGGAGCTCACAACGACACGC 221
 QY 63 AGCCCGAGCAGCCG--GGGGGAGGACACCGCGAGCGAGCTGTGCTCAGTGGCAT 119
 Db 222 AGCCCGAGCAGCCGCGGAGGGAGCTCCACCGCGCTAGCCCGAGCCGCGGCGC 281
 QY 120 GGGCCCGAGCTGAAGCTTGGGA-----GCACGTGAGACCCCGCCGAGCGGAC 170
 Db 282 GGGCCCTCGGCGGAGGCGGCGCCAGACACACCGCGGACCCCGCATGCTGCTCGAC 341
 QY 171 CCGCGCACCAAGCTCCACAGAAAGATGGCAGCTGCTTCTGTCAACGCGGTAGTGA 230
 Db 342 CCGCGCACCAAGCTCTACAGAAAGATGGTCACTGCTCCACATCAATGGGTAGCTGAG 401
 QY 231 CAAGAGATGTCATGTCACAGAGAAAAACGAGAGG----- 268
 Db 402 CAAGATGAGCTCAGCGCTCAGAGAGGTGACCTTAATGCGCAAGAAAGAGCCCTGAACGCT 461
 QY 269 -----GCAGAGAAAGATCGTTGATGAGATGTTGACAGCA 308
 Db 462 CAAGAGCCCTTAAACAGCCAGAGAGAAAGAAAGTCAATTGTCAGCGAGTTGACAGAGA 521
 QY 309 GAGTCAGAAAGTGAAGAGAAAAACCGAGTTGAAAGAAATGGCGCCAACTCCACAGCT 368
 Db 522 GACTTGAAGATGTAAGAGAAAGATCTCCGATTAAGATGGCTACTTAATCAGCGTT 581
 QY 369 GTTGAAGATATCAGAAAGATGGCGAGAGAGACATCAGAAATATTTGAACAGATCCT 428
 Db 582 GTTCACGACATCAGATGATGGGAGAGAGAGA--ACGAAATATGAAACAGATTCCT 638
 QY 429 GCTTGAAGAAACATGTGGAAGAAATGTAACGCTGCTGCTCCGAGCTTAATGATGTT 488
 Db 639 TCTTCAGAAACCAATTTAAGAGAGCTAACACACCTCAGTGAAGCTTAATGATGTT 698
 QY 489 GGCCTCAAGAAAGATTTAAATTTGTTTAAATTCAGCGTGAAGAGATTTAAAT 548
 Db 699 GATTTAAAGAAAGTGTTTAAGTTGTTGCTTTAAATTCATCTGTAAAGAAAGATTAAGCA 758
 QY 549 GAAAGTCAGATCTGTCAACTCACTGTCAAGAGATGAAGGAGGAGGCGCAGAA 608
 Db 759 GAGAAAGCTGACACTGTCCAGTACTCACTGTGAAGAAAGATGAAGGAGGAGCAGCA 818
 QY 609 GCCTGTGCGAGCTGGAGACCAACAGAGCCCACTGTGAGACTGCGCTGGAGAGTCA 668
 Db 819 -----GGGGCTGGCGACCAACAGAGCCCACTGTGGGCT-----GAGAAACCA 863
 QY 669 GCATCCAAAGAAAGTGAAGCTGAGCAATCCAAGAGAAAGCAAGAGCACCTGAAGCA 728
 Db 864 GCATCCAAAGAAAGGAAACCCAAATCTACAGAAACCCGAAAGAACCTTGAAGCT 923
 QY 729 GAACAGAGCAGCAGAAATCCCTTCAGCGCAATCTGATCAAGCGCTGAGAGAA 788
 Db 924 GAGCAAAAGCAGCAGAAATTTCTCCCGAGCGAATCTGGCGCAAGCAGT--GAGAGAA 980
 QY 789 GCCAAAGATGAAGGAAAGAAAGAAAGAGAGAGCCCACTAATCCCAAGATTCCTCCG 848
 Db 981 TGCAAAGAGGAGGAGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1040
 QY 849 ACCAGCCAGTCAAGTGAAGCAATCTTCTCAAGAAAGTTCTTCACTCAGCGTTGG 908
 Db 1041 ACTAATCCGTACAGATGAAGAAAGAGATCACTTCAAAAATTTTCACTCAAGGTTGG 1100
 QY 909 GCGGCTGCGCAGAAAGCAGCTTCAAGAAATCAAAAGAGATGATCTGAAAACTGCC 968
 Db 1101 GCGGCTGCGCAGAAAGCAGCTTCAAGAAAGCAGAAAGAGATGAAGTGAAGCTTCA 1160
 QY 969 GAGAAAGAAAG 1028
 Db 1161 GAGAAAGAAAG 1220

QY 1029 CCAGCCTCGAGGA-----GCAGGAGCCGGCAGAGAC 1061
Db 1221 GTTGCCCTCCGAGAACTGACCGCCTCCGAGCAAGCCACCACACAGGAGCCGGCAGAAAGT 1280
QY 1062 ACAGACAGGCCAGGTTGTTCAGCAGACTACGAGAAGGTGAGCTGCCCTTTGGAAGACCAG 1121
Db 1281 GCCACAGCCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGGACCA 1340
QY 1122 GTTGGTGACCTGAGGCGATCGTCAGAGGAGAGTGTCTCTCTTTGGCAACGGAAGTGT 1181
Db 1341 GTCAGTGGCTCGCAGGACCTCTCTGAAGAGAAACCTCTCGTTGGCGACAGAAGTGT 1400
QY 1182 GATTGAGAAGATGGAAGCCCAACCAAGRA--GTTGTTCCAGAGTCCACGTCGACACCGTG 1238
Db 1401 GATGAGAAAATAGAAATGACCAAGAAAGAGGTTGTGCGCGAAAGTCCACGTCAGACCCGTG 1460
QY 1239 GAGAAGACAGAGGAGGAGCAGGAGGAGGAGAGGAGCTGAAAGGGGCGTGGTGGTAGAA 1298
Db 1461 GAGGAGAAACCGAAGACAGAAA-----ACGGAGGTGGA 1496
QY 1299 GGAACAGGAGAAATCCTTGGCCCCCTGAGAAACTGGCTGAGCCCCCAGGAGGTCCCCCAGGA 1358
Db 1497 GAAACAGCAGGGTCTGTGCCAGCTGAAAGAAATGGTTTGGAAATGGATGCAGAAACCTCAGGA 1556
QY 1359 GCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTCGAGGAGACCAC 1418
Db 1557 GCCGAACCTGCCAGGAGCTGTGGAAGCTCAAGAAACGTTGTGTTCGGAGAGGACCT 1616
QY 1419 ACTCACTGACAGACCTAAAGTCTCTGAAGAGAAAGCGCTGCCAAACACCCAGAGGCA 1478
Db 1617 ACACGGAGCTGACCTCAGTCTGATGAGAAAGTGCTGTCTCAAACCCCGAAGGGCT 1676
QY 1479 GTCAGTGAAGTGAAGTGTCTCTCAGGAAAGAAATCAAGGTACAGGAAGTCCCTTG 1538
Db 1677 GTGAGTGAGGTGGAATGTCTCATCACAGAGAGAAATGAAGTGCAGGGAAGTCCACTA 1736
QY 1539 AAGAACTCTTCAGTAGCTCAGGCTTAAAGAGCTGTCTGGAAAGACGAGAGGGGAAA 1598
Db 1737 AAGAAGCTTTTACCAGCACTGGCTTAAAAAAGCTTCTGGAAAGAAACAGAAAGGAAA 1796
QY 1599 CGAGGAGTGGGGAGACGAAGAGCCTTGGAGAAATACCAACACATTTCACCGAATCCCCA 1658
Db 1797 AGAGA---GGAGGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCGATTCCTCG 1853
QY 1659 GAGAGTCTGATGAGCAGAAAGGAGAGAGCTCTGCGTCTGCCCGAGGAGCCTTGAGGAG 1718
Db 1854 GACAGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCCTCATCCCTCTGAGGAGCCCGAGGAG 1913
QY 1719 ACCACGTGTCTGGAGAAAGGCGCTGGAAGCACCCAGAGTGGGGAAGCTGAGGAAGGA 1778
Db 1914 ATCACGTGTCTGAAAAGGGCTTAGCCGAGGTGCAGCAGGATGGGGAAGCTGAAGAGGA 1973
QY 1779 ACTACTTCCGATGAGAGAAAGAGAGAGAGGATCACTCCCTGGGCATCCTTCMAAAG 1838
Db 1974 GCTACTTCCGATGGAGAGAAAAGAGAAAGGTGTCACTCCCTGGGCATCATTCMAAAG 2033
QY 1839 ATGGTGAACCCCAAGAACCGGTCGGAAGACCTTCTGAGAGTGACAGAGGAGGAAGAGCTG 1898
Db 2034 ATGGTGAACCCCAAGAGCGTGTAGACGCTTCGGAAGTGTATAAGAAAGATGAGCTG 2093
QY 1899 GAGAAGTCAAGAGCGCCACTTGTCTCCTCATGATAGCAAGTGTCAAGAAATGCAAGAT 1958
Db 2094 GACAAGTCAAGAGCGCTACTGTCTTCCACCGAGAGCACAGCCTCTGAAATGCAAGAA 2153
QY 1959 GAAGTCAAAACTGTTGTGAGGACAAAGCCAGAGAAACCAAGCGTAGGTGGATACT 2018
Db 2154 GAAATGAAGGGAGCGTGGGAAGAGCCAAAGCCGGAAGAACCAAGCGCAAGGTGGATACC 2213
QY 2019 TCAGTGTCTTGGGAAGCACTGATTTGTGTGGATCATCTCCAAAGAGAGAGCAAGGAAGCA 2078
Db 2214 TCAGTATCTTGGNAGCTTAATTTGTGTGGATCATCCAAAGAAAGAGCAGAGGAGG 2273
QY 2079 TCCTCTTCAGATGATGAAGGAGGCGCAAGGACACTGGGAGGGGACAGTCAAGAGCAGAG 2138

Db 2274 TCCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGAGGAGACCAACCAAGAAAGCTGAT 2333
QY 2139 GAGGCCAGCAAAAGCAAAAGAGCCGGAACAGACGCTGTTCTCTCCAGCAGCCACAGGAGCAG 2198
Db 2334 GAGCCCGGAAAAGACAAAGAGACGGGACAGACGGGATCCTTGCTGTTCCCAAGAACAT 2393
QY 2199 GACCAAGCGCAAGAAAGTTCTCTACCCGAGCAGCGGGAAGCCCTTCCGAAGGGGAAGGT 2258
Db 2394 GATCCAGGGCAGGAAAGTTCTCTCCCGAGCAAGCTGGAAGCCCTACCGAAGGGGAGGC 2453
QY 2259 GTCTCCACCTTGGAGTCATTAAAGATAGTCACCTCCAGAAAAAATCCAGTCAAAA 2318
Db 2454 GTTTCACCTTGGAGTCATTAAAGAGTTAGTCACGCCAAGAAAAAATCAAAAGTCCAAG 2513
QY 2319 CTGGAAGAGAAAACCGGAAGAC-----TCTAGTGTAGAGCAGTTGTCCACTGAG 2366
Db 2514 CTGGAAGAGAAAACCGAGACTCCATAGCTGGGTCTGGTGTAGAACATTCCTCACTCCAGAC 2573
QY 2367 ATCGAACCGAGTAGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCCGAGCGCG 2426
Db 2574 ACTGAACCCCGTAAAGAGAAATCTCGGTCTCAATCAAGAAAGTTTATTCCTGACGAAGG 2633
QY 2427 AAGAAAGGCGCAGACGGGAGCAAGAACAGCCACTGTGGAAGACTCAGGCGCAGTGGAG 2486
Db 2634 AAGAAAGGCGCAGATGGGAAACAAAGAACAGCCCTGTTGAAGACGAGGGCCCAACAGGG 2693
QY 2487 ATAAATGAGGACGACCCCTAATGTCCAGCGCTGCTCTGTCTGATATATATGCAGTG 2546
Db 2694 GCCAAACGAAGATGACTCTGATGTCCCGCGCTGCTCTGTCTGATATATGCTGTA 2753
QY 2547 GAGAGGGAGAAAGATGG-----AAGCCAGGGGAATACGGAGTGCCTCCAGCTGCTG 2597
Db 2754 GAAAGGAGAAAATGGAGGCACAGCAAGCCCAAAAAGGGCGACAGCAGCCCGAGCAGAAG 2813
QY 2598 GGGCTGTGTACGTGTCCGAGGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTCCGA 2657
Db 2814 GCAGCCACTGAGGTGTCCAAGGAGCTCAGCGAGAGTCAGGTTTCATATGATGGCAGCAGCT 2873
QY 2658 GTCATTGTATGGGACACGAGGCGAGTCACAGTGTCTGAAGAGCGGTCTCTTCTGGTATATCC 2717
Db 2874 GTCGCTGACGGGACGAGGCGAGCTACCATTTATGAAGAAAGGTCTCTTCTTGGATATCT 2933
QY 2718 GCTTCGCTAACAGAACTCTTTGAACACACAGCGGGAGAGCCATGCCACTGTTGAAGAG 2777
Db 2934 GCTTCAGTGAAGAACTCTTTGAACAGTAGAAGCTGAAGCCGCACTGTTAACTGAGGAG 2993
QY 2778 GTCACCTGAAAAGACATCATTC-----AGAAAGAACTCTGTGCTCACCAGAGCTTACCA 2834
Db 2994 GTATTGGAAAGAGAAAGTAATTGCAAGAAAGAAACCCCCCAACGGTTACTGAACCTCTGCCA 3053
QY 2835 GAGGTTAAAGATGCCCATGACGACATGGTCAACAGTGAAGTGGATTTCACTCAGAAAGCT 2894
Db 3054 GAGAACAGAGAGCCCGGGGCGACACGGTGTGATGAGCGGAATTTGACCCCCGAAGCT 3113
QY 2895 GTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTCTGAAGAAAGTTACCGAAGCATCGGG 2954
Db 3114 GTGACAGCTGCAGAAACTGCAGGCGCATTTGGGTTCCGAAGAAAGAACCGAAGCATCTGCT 3173
QY 2955 GCCGAAGAGACACACAGACATGTTGTCGAGATTTCCAGCTGACTGACTCCCCAGACACC 3014
Db 3174 GCTGAAGAGACCAACAGAAATGGTGTGAGGAGTCTCCAGTCTCCAGTAAACCGACTCCCCAGACCC 3233
QY 3015 ACAGAGGAAGCCACCCAGTTCAAGGAGGTAGAGAGTGTGCTAGATACAGAGAGAGAG 3074
Db 3234 ACAGAGGCGCCACTCCGTTGAGGAGGTGGAAGGTGGCGTACTCTGACATAGAAAGAGCAA 3293
QY 3075 GAGCGCCAGACGAGGCCCATCTCTCCAAAGCGTTGCAGACAAGGTGAAGAGGAGTCCAG 3134
Db 3294 GAGAGCGGACTCAAGAGTCTCTCAGGCGAGTGGCAGAAAAAGTGAAGAGGAAATCCCAG 3353
QY 3135 GTGCCCTGCAACC---CAGACTGTGTCAGAAACGGGGTCAAAAGCACTGGAGAAAGTTGAG 3191

3354 CTGCTGGCACCGGCGCCAGAAAGTGTGCTTACGCTGTGACAGAGACAGAGGACAGA 3413
QY 3192 GAGGTAGAGAGAGACTCCGAGTGTGCTTCCGAGAAAGAGAGACGTTATGCCGAAA 3251
Db 3414 AGACAGAGAGAGAGGCTGAAAGCGTCGGGTCTGAAAGAAAGACGAGTAGTGTGAAA 3473
QY 3252 GGAACCCGTGAGAGAGCTGAGCTGAGCATCTTGGACAGAGGCTGTGACAGTGGACAGCT 3311
Db 3474 GTAGATGCTCAGAGAGGCAAAAATGAGCTTTTACACAGAGAGAGAGGTGTGGGACACC 3533
QY 3312 ACTCAGAGAGAGCTTGG---AAGTCTCTGAAGTCAAGGAGATGAGAC-----CAT 3359
Db 3534 ACCCAGAGAGAGCTTGAAGAAAGCTCTCAAGTCAACAGAGAGATGAGTCCAGTACCTT 3593
QY 3360 GTCCGACAGTCCCA-----GGTTATCAGCTCCAGCAGCTGATGGA 3401
Db 3594 GTAAACCATTTGTCAAGCCGAAAACCTTAGCTGGGGTAAATCAACAGAGATGTGTGGA 3653
QY 3402 CAGGCGCTGGCCCCCTGAGTCACTCCGAAACCTTGACAGACAGAGACAAATGGAACACT 3461
Db 3654 CAGGCTATCCCTCCCTGACTCGGTGGAACCTCTACAGACAGTGAATGTATGGAAGCAC 3713
QY 3462 CCTTAGCAGATTCTGACACTGCTCAGATGGGACACAGCAAGATGAAACCATTGACAGCCAG 3521
Db 3714 CCGGATCCGACTTTGAGACGACACAGGACAAACCCAGAAAGAGAGATTTGTGAAATCCAT 3773
QY 3522 GACAGTAAAGCCACTGACGCTGTACAGGAGTCAAGGTCAAGAAAGAGAGGCGCTACT 3581
Db 3774 GAGGAGAAATGA---GATGATCTGGTACAGTCAAGGGGACAGAAAGCAGAGCAATTCCT 3830
QY 3582 GCTCGAAAGAGAGAGCTTCCGACACTACTAATATTTCCAGCCCAAGAAAGATGAG 3641
Db 3831 GCACGAAAGAGAGAGGCTTCCAGACCTTCAAGTTTGTTCAGAGAAACCTAAAGAA 3890
QY 3642 GAAAGACCGAGAGAGATTTCTTGAACCTACACAGCAAGACCTTACTGCTCAGCCGTG 3701
Db 3891 CAATCAAGATGGAAGACACTCTAGAGCATACAGATTAAGAGGTGTCACTGGAACCTGTA 3950
QY 3702 CCGGTTCTGGCAAAAGACTGAGGTGGGTCAAGAGGTGA-----GGTTGACTGTGG 3752
Db 3951 TCCATTCTGTCAAGACTGAGGGGACTCAAGAGGCTGACAGTATGCTATGAGAAAC 4010
QY 3753 GATGAGAGAAAGTCAAGAAAGACAGAGGTGTGTACTCTGG-----ACCCAC 3806
Db 4011 AAAGCATACCAATTTTCCAAAGACTTGAAGGGTCTATAGACACAGGCATTAACGTCACT 4070
QY 3807 AGTCAAAAGGCTGCTGATGTGACATATGACATGAGTATGAGGTGCGCGGTGTCCAG 3866
Db 4071 CGGGAAGAGTCACTGAAAGTCCCTTAAAGGTGAAGGAGACAGAAAGTGAATGTAAA 4130
QY 3867 GAAAAGAGAGTACTGAAGTGCAGAG-----TCTTAAGCTTGAAGAGGA 3911
Db 4131 AAGATGATGCTCTGAACTGCAGAGTCAAGCTAAGTCTCTCCCTCCCTGGAGAGA 4190
QY 3912 GAGATGAGAACTGACGTTGAAAGAGAAAGGAGACAAACCCAGAGAAATGTAGTGA 3971
Db 4191 GAGATGTGTGTTCAAGTCCAAAGGAGAAACAGAGACAGACCAACCATGTGATGTA 4250
QY 3972 GAAG---TGACAGAGAAACAGCGCTCTGAGCATGAAGAACTTACGGGAAGCCAGT 4028
Db 4251 GGAAGCTTGAAGCAAGAAAGCTGTTACCGTATCTGAAGAGTCAAGTAAAGCTCTC 4310
QY 4029 CTGACACTTGACATGCCAGCTCAGAGAGGGGAAAGGCACTGGGAAGCCTTGAAGGAAC 4088
Db 4311 CAGACAGTGAATGTGCCCTCATATAGATGGGCAAGAGAGTCAAGCTTGGAAAGAAC 4370
QY 4089 CCTTCTC---TCCAGACCAAGCAAGAGGTTGATGAGAGTTAAAGTCTAAAGCTCG 4145
Db 4371 CCTCTCCCTGCTAGGTCAAGGAGGAGCATATGACCAAAATTCAAATCTCAGACTCT 4430
QY 4146 GACCAACAGCTCAACCAAGCAAGAGAGCTGTGAAAA-----GGTCAATGAAAAAGTT 4199
Db 4431 GAGGATCATCTCACTTAACAGCGGCTGCAGAGAGAGAAAGTCTTAAAGAGAACTGCC 4490

QY 4200 GTGATTTCAAGACAGGTGAAGTCCAGAGGTGTAGTGACACTTATTAACAGCTGAG 4259
Db 4491 AACATTTTGAAGACAGGTGAAGCTTGAAGCTGCAGAGTGACATTTAGTTCTGGAAGAG 4550
QY 4260 AAGTCTCTGCAAGGGGTGGCACTGAGACTTTCAGCATGACAGAGACAGCGTACCCCTG 4319
Db 4551 AATCTCTGAAAAAATGAAAGCTTTGGCCGCTCATCCAGGGGAAGATGTGTGCCACA 4610
QY 4320 GGGCTGAGTCTCAGGACAGAAATCCATCCATCATATGTAATCTCTGCTCTGAAAGCAC 4379
Db 4611 GGGCCGACGTCTCAGGCAAAATCGACACAGTATATCTGTACTTACCAAAAAAGGC 4670
QY 4380 CTATATCCTGACCTACAAAGAGAAATTAAGGCATCCACAGAGAGAGCATCAGAGAGAG 4439
Db 4671 TTAAATTCGACCTGGAAGAGAGAAACACATCATCTGAATGGAAGTCAAGATGAATGC 4730
QY 4440 GACAAAGCAGATGCTGTGCTGATGCTGACCGCAAGAGAGTACAGCAATCGAAAAAGTC 4499
Db 4731 GATAGCAGAGTTGCTTGCC-----AGAGGTCAAAAGTGAATGATGACATTTG---AGGAT 4781
QY 4500 CTAAAGCTGAACCTGAGATCTTGAACCTTGAGATGAAGCAACAAATTTGCTGAAC 4559
Db 4782 TTAGAGCTGAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTTAACTGTCCAAAAC 4841
QY 4560 GTCAATTCAGACAGCCGTTGACCAATTCGACAGTAC---AGAAAGCCCGGAAGCAT 4616
Db 4842 ATCATTCAGACAGCCGTTGACCAATTTGTACGTAACAGAAAGACGCCGAAGATGTTG 4901
QY 4617 GCTTATGATTCACAGACCAGTTC-----CTGCATGACAGCTTGAACAG 4660
Db 4902 ACGTGTAGTTACAGACCAAGCTCAGCTGATTAAGCTGACACCGAGAGCTGTGACAG 4961
QY 4661 CAGGAGCCCAACAGATCTGGAACAAAATGAATGCCAAGTGAACACCAAGTCCC 4720
Db 4962 GAAACGAGAAAGAGAGAGAGAACTCAGGCTCTGCACAGATGAACACCAATTA 5021
QY 4721 GCAGCCCA-----GAGAGACTTGAAGTCCGACCGTTCTGGA 4759
Db 5022 TCAGCAAAAGAGAGTCAAGTCAACCGCAAGTGAACAGCAATTCGATATTTCCAAA 5081
QY 4760 GGCATGGGCTCAGCCTCGGAATGCTTGGCCGCTTGCAGTTGAAGGCGCGGTGCA- 4818
Db 5082 GACATGATGAAGCTTCGAAAGAACATGACTGTGAGATGAAGATTCACATGTAAAT 5141
QY 4819 ----- 4818
Db 5142 GATCAGAGCTGGAAGAGTGTCTCCATCTGAGAAAGAGAGGTGAGCTGGAACA 5201
QY 4819 -----AGTAAGATTAAGAGTGTCT 4840
Db 5202 AAGTGTGCAAGATGATGATGATGCTTGTGAGCAAGAAATGAAAGTCACTA 5261
QY 4841 CCTCAACCAAGA--TCCAAAGAGAGATGCTGTGAGCTCTCAGCTCAAAAGCTTA 4898
Db 5262 GTTGAACCGAAAGAGATGAAGAAAGGTGATGTGTGATGACCTGAAACCAAGACTCA 5321
QY 4899 GCCCAGGAGAGGCCAGTCTTGGAAAACCTAACCAAGAAATCCAGACCAACGGA 4958
Db 5322 GCCCTGGCTGATGATGATGCTCAGAGAGGCTTAAACCAAGAGTCCCGCATTAACAAATGGA 5381
QY 4959 CCAAGCTTAACCGAGAGAGGCGATCCCCAAAAGTTGAGGTTCAGAGAAAGAAATGTCT 5018
Db 5382 CCAAAACAAAAGAGAAAGAGAGATGCCAGAGAAATGAAATTTGCAAGAGAGAAAGTGCAC 5441
QY 5019 ACCAAGTCACTCAAGAGAGAACAGGCCAGGACAGAGAGAGACCTGACAGAGCAAAAGGA 5078
Db 5442 AGTGAATCAATTAAGCATCAACCCCAAGACAGAGAGATTAACAGAAACAAAGAGA 5501
QY 5079 GACCTGGCAGATCTTAAGATTTAGTTG-----CTCATTTGTAC 5117
Db 5502 GAATCTGAAAGTCAAGACTTAACAGAACTTAAACATCATGAGTTAACTCATTTGTCT 5561

Qy	5118	ATCTGTAAGACCAAGATGTGAAAAACAAGTCA CAGAACAA ----GATGCTGCTGTGGGAC	5173
Db	5562	GTTTGGGAAGCAGATGTGAAACACAAGTAGTAGAAGAAAATGAATGCTGCTGTGAGA-	5620
Qy	5174	CTTCAGACCAAGATTTTCAGAGCCATCAGATCCAGAGACGAGGCGCGTCCAATCATTTCC	5233
Db	5621	CTGAAGACCAAGTATTTCAG- ACTTTGGAATTTGGAGAGCAGGCACATCAACTGATCTCA	5679
Qy	5234	ACCCAGTAGACACCCCGCAATTTCTGAGGCTTTCATCCGGAGCTAGAGCCAGCTAAACATT	5293
Db	5680	TTTCTAGAGAGC-CCCTGACAACTCTGAGCTTCATCAGAGCTAGAGCCATTTAAACATT	5738
Qy	5294	TCCTCGTTTCAAGACTGCTTGAATTTGGTCCCTTTGATGCGGTTCGATTTCTTAAACATT	5353
Db	5739	TCCTCTTTTCCAAGACCAACCTACAAATTTTCCCTTGATAAC-----	5778
Qy	5354	TCCTCGTTTCAAGACTGCCTTTGATTTGCCCTTGTATGCCGTCCGTGATTTTCGGATTTA	5413
Db	5779	-----CATATAAAATCTGATTTTA	5796
Qy	5414	AGGTCCTGCGTTTCTCAACCTGGAAC-CAATTTGCCATACCTAGTTTCCACATTTCTCAAACT	5472
Db	5797	AGGTCCCTAAATTTCTTAACCTGGAACGTGGAGTTGGCAATACCTAGTTCTGCTTCTGAACT	5856
Qy	5473	GGAGCATCTCTTTATGATTTATATATGATGTTTATGTAGTCTCTCTCTGTACCTAT	5532
Db	5857	GGAGTATCATTTCTTACATATTTATATGATGTTTAAAGTAG---TCCTCTGTATCTAT	5913
Qy	5533	TGTTATATTTTTTTCTAACGTTT-----AGGCATGCTTTTTTGTTATTA	5575
Db	5914	TGTTATATTTTTTTCTTAATGTTTAAAGAAATGTGCAGGATACTACATGCTTTTGTATCA	5973
Qy	5576	TGCAATATATAACGGGTGTGCAGCCATAGCAGCGTTTGAAGAAAGCTCCAGCCCTCAACTG	5635
Db	5974	CACAGTATATGATGGGCGATGTGCCATAGTCAGGCTTGGGGAGCTTTAAGCCTCAGTTA	6033
Qy	5636	TAACTCGAGCAACAGATAACATTC-----CTGGCAAGAGAGACAAAGTCTTTTT	5686
Db	6034	TATAACCCACAAAAACAGAGCGCTCCTAGATGTAAACATTTCTGTCAAGGTACAAATTCCT	6093
Qy	5687	TAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGTGGTGTGTTTCCT	5746
Db	6094	TAAATTCATTAATGATGAGTCCATATTTAGTGGTACTCTGAAATTTGGTCACTTTTCT	6153
Qy	5747	ATGCACAGCGAGCTCAGAAATMAAACCCCATTTTGAACACATCAGAGATGCCCAATATT	5806
Db	6154	ATTACCGAGTGTGCCAAAACCTAAAGAACATTTGAAGACATACAGATGTTCTATTGTC	6213
Qy	5807	ACCAATATTTTTTCCCGCTTTTGTGTAATCCAGTCCAGGTGGAAAGAAGTCTCCTCTG	5866
Db	6214	ATTGGGAAATTTTGC-----TTTCTAACCCAGTGGAGGTAGAAAGAAGTTATATTCT	6266
Qy	5867	TGTCAGA-----TTAAGCCCTGTCTTAAATGATATGGAATAATGAGT	5909
Db	6267	GGTAGCAAAATTAACCTTACATCCTTTTCTACTTGTATTGTTGTTGGACCGATAAGT	6326
Qy	5910	GTGCCTAAGCCCATGAGA-----TGTTTCTTAATGCAGAAAGGATCTGT	5953
Db	6327	GTGCTTAATCCTGAGCAAGTAGTGAATATGTTTATATGTTATGAAGAAAGAATGTT	6386
Qy	5954	TGTACGTTTTTTTGAATGTACTCTTCTATGTGGACCGAAATTCATATGCAGATCGAAAGT	6013
Db	6387	TGTAAG- TTTTTTGATTCTACTCTTATATGCTGGACTGCATTCACATGSCATGAAATA	6444
Qy	6014	AGTCCTGTTCTTTACAGATGGTATTTTGTATGATGACTGGAGTTTGTCTGTGTTATATCTG	6073
Db	6445	AGTCAGGTCTTTTACAAATGGTATTTTGTATAGTACTGGAATGTTGTGTGCCATATTTG	6504
Qy	6074	TGCCCTTCTTTAAGAACAAATGTCATTATCTCTTCTTTGGATAAATTTGTGATTTGACAA	6133
Db	6505	TGCCATTCCTTTAAGAACAAATGTTGCAACAACTCATTTGGATAGTTGTGATTTGACGA	6564
Qy	6134	CTGATTTTAAATAAA 6147	

Db 282 GGGCCCTGCGCGAGGCGCGCCAGACACACCGCGGACCCCGCATGCTGCTCCGAC 341
Qy 171 CCCCACCAAGCTCCACAGAAATGGCAGCTGCTTGTGTACACGCGTAGCTAA 230
Db 342 CCGGCAACCAAGCTCTACAGAAATGCTCAGCTGTCCACATCAATGGCGTAGCTAG 401
Qy 231 CAAGAGATGTCATGTCCAAAGAGAAAACGAGAGG----- 268
Db 402 CAAGATGAGCTCAGGCTCCAGAGGGTGACCTAAATGCGCAGAAAAGACCTGAACGCT 461
Qy 269 -----GCAGAGAGAAAGTCTGTGATGAGAAATGTTGACAGCGA 308
Db 462 CAAGAGCCCTAAACAGCAGAGAGAAAGAAATGCTATTTGTACGAGAGTTGACAGAG 521
Qy 309 GAGTCAGAAATGTGAGAGAAAAAAGCCGAGTTGAAGAAATGGCGCACTCCACAGCT 368
Db 522 GACTGTGAAGATGTGAGGAGAAAGACCTCCGATTAAGAGATGGCTACTAAGTCAGCGGTT 581
Qy 369 GTTGAAGATATCAAAAGATGGCAGAGAGACATCAGAAATTAATGAAACAGATCCCT 428
Db 582 GTTCAAGCATCAAGATGATGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 638
Qy 429 GCTTGAAGAAACAATGTGAGAGAAATGGTACAGCTGAGTCCAGGCTAATGATGTT 488
Db 639 TCTTGAAGAAACAATTTGAGAGAGCTAACACACCACTGAGTCCAGGCTAATGATATT 698
Qy 489 GGCCTCAAGAAAGTATTTAAATTTGTGTTTAAATTCACGCTGAAGAGATAAAT 548
Db 699 GGATTTAAGAAAGTGTTTAATGTTGTGCTTTAAATTCAGTGAAGAAAGATTAAGACA 758
Qy 549 GAAAGTCAGATCTGTCCAATCTACTGTCAAGAGAGATGAAGGCGAAGGCGAGAA 608
Db 759 GAGAGCCCTGACACTGTCCAGTACTCACTGTGAAGAAAGATGAAGGCGAGAGAGACA 818
Qy 609 GCCTGTGTGAGCTGGAGACCAAGAGAGCCAGTGTGAAGTCCGCTCGAGAGTCA 668
Db 819 -----GGGGCTGGCGACCAACAGAGCCCAAGCCTTGGGGCT-----GAGAGAGCA 863
Qy 669 GCATCCAAAGAAAGTAGCTGAGCAATCCACAGAGAACAGAAAGCAACCTGAAAGCA 728
Db 864 GCATCCAAAGAAAGTAGGAAACCCAAACATCTACAGAGAAACCGAAGAGACCTGAAAGCGT 923
Qy 729 GAAACAGAGCAGACAGAAATCCCTCTTCAAGCCGAATCTGATCAAGCGCTGAGAGAA 788
Db 924 GAGCAAGGCAAGAGAAATTTCTCCCGAGCGAATCTGGCCCAAGCAGT---GGAGGAA 980
Qy 789 GCCAAAGTAGAGAGAGAAAGAAACAGAGAAAGAGCCCAAGTCCCGAGATCCCG 848
Db 981 TGCAG 1040
Qy 849 AGCAGCCAGTCAGACAGTGAACAACATCTTCTTCAAGAAATCTTCACTACAGGTTG 908
Db 1041 ACTAGTCCCGTGAACAGTGAACAGATCAACCTTCAAAAAATTTCTTCACTCAAGGTTG 1100
Qy 909 GCGGCTGCGCAGAGAGACCAAGCTTCAAGAAATCAAAAGAGATATCTGAAACTGCC 968
Db 1101 GCGGCTGCGCAGAGAGACCAAGTTTCAAGAAAGCCGAAAGAGATGAATGGAAGCTTCA 1160
Qy 969 GAGAAAGAGAAAG 1028
Db 1161 GAGAAAGAGAAAG 1220
Qy 1029 CCAGCCTCGAG 1061
Db 1221 GTTGCCTCCGAG 1280
Qy 1062 ACAGACAG 1121
Db 1281 GCGCAG 1340
Qy 1122 GTTGTGAGCTGAGAGAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1181

Db 1341 GTCAGTGGCTCGCAGGAGACCTTCTGAAGAGAAACTGCTCCGTGGCGAGAGAGTGT 1400
Qy 1182 GATGAGAGATGAG 1238
Db 1401 GATGAGAGAGATGAG 1460
Qy 1239 GAG 1298
Db 1461 GAG 1496
Qy 1299 GGAACAGAGAGATCTTCCCTCGAGAGAACTGCTGAGAGAGAGAGAGAGAGAGAGAG 1358
Db 1497 GAAACAGAGAGAGTGTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1556
Qy 1359 GCTAGCCTGTGAG 1418
Db 1557 GCCAGACCTGCAAG 1616
Qy 1419 ACTCAACTGACAGACCTAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1478
Db 1617 ACAG 1676
Qy 1479 GTCAGTGAAGTGAAGATGCTGTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1538
Db 1677 GTGAGTGAAGTGAAGATGCTGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1736
Qy 1539 AAGAACTCTTCAAGTCTCAG 1598
Db 1737 AAGAGCTTTTTCACAG 1796
Qy 1599 CGAGAGAGTGGGAG 1658
Db 1797 AG 1853
Qy 1659 GAGAGTGTGATGAGCAG 1718
Db 1854 GAGAGCCAG 1913
Qy 1719 ACCAGTGTCTGAG 1778
Db 1914 ATCAGTGTCTGAG 1973
Qy 1779 ACTACTTCGATGAG 1838
Db 1974 GCTACTTCGATGAG 2033
Qy 1839 ATGTGACACCAAG 1898
Db 2034 ATGTGACGCGCAAG 2093
Qy 1899 GAGAGAGTCAAG 1958
Db 2094 GACAGAGTCAAG 2153
Qy 1959 GAAATCAAAACCTGTTGTGAG 2018
Db 2154 GAAATCAAAAG 2213
Qy 2019 TCAATGTCTTGGAG 2078
Db 2214 TCAATGTCTTGGAG 2273
Qy 2079 TCCTCTTCAATGATGAG 2138
Db 2274 TCCTCTTCAATGATGAG 2333
Qy 2139 GAGGCGACAG 2198
Db 2334 GAGGCGGAG 2393
Qy 2199 GACCAAGCGCAAG 2258
Db 2394 GATCCAGGCGAG 2453

Qy	2259	GTCTCCACTTTGGGAGTCATTTAAAGATTAGTCACTCCAGAAAAAAATTCOAAGTCAAAA	2318
Db	2454	GTTTTCCACCTCTGGGAGTCATTTAAAGGGTTAGTCA CGCCAGAAAAAAATCAAAGTCCAG	2513
Qy	2319	CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTGTCTCCACTGAG	2366
Db	2514	CTGGAAGAGAAAGCGNAGNCTCCATAGCTGGGTCTGGTGTAGAACATTCCACTCCAGAC	2573
Qy	2367	ATCGAACCGGATGAGAGAAATCTTGGGTTTCCATTAAAGAAATTCATCCCGGAGCGCGG	2426
Db	2574	ACTGAACCCCGTTAAAGAGAAATCCTGGGTCTCAATCAAGAAAGTTATTCTCTGGACGAAG	2633
Qy	2427	AAGAAAAGGCGACGCGGAGCAAGAACCACTGTGGAGACTCAGGCGCAGTGGAG	2486
Db	2634	AAGAAAAGGCGCAGTGGAAAAAAGAAACAAAGCCCTGTTTGAAGACGCAAGGCGCAACAGGG	2693
Qy	2487	ATAAATGAGGACGACCCTAATGTCCAGCGCTCGTGCCCTCTGTCTGAGTATTAATGSCAGTG	2546
Db	2694	GCCAAACGAGATGACTCTGATGCTCGCGCGGTGGTCCCTCTGTCTGAGTATGATGCTGTA	2753
Qy	2547	GAGAGGAGAAAGATGG-----AAGCCACAGGGGAATA CGGAGCTGCCCCAGCTGCTG	2597
Db	2754	GAAAGGAGAAAAATGGAGGCACAGCAAGCCCAAAAAGGCGCAGAGCAGCCCGAGCAAG	2813
Qy	2598	GGGCTGTGTACGTGTCCGAGGAGCTCAGTAAAGCTCTGTGTCCACACTGTGAGTGTGCGA	2657
Db	2814	GCAGCCTAGTGTGCCAAGGAGCTCAGGAGAGTCAAGTTTCATATGATGGCAGCAGCT	2873
Qy	2658	GTCAATTGATGGGACGAGGCGAGTCACCACTGTGCAAGAGCGGTCTCTCTCGTGGATATCC	2717
Db	2874	GTGCTCTGACGGGACGAGGCGAGCTACCATTAATTGAAGAAAGGTCTCTCTCTTGGATATCT	2933
Qy	2718	GCTTCGTATAAGAACTCTTTGAACACACAGCGGGAGAGCCATGCGCACCTGTTGAAGAG	2777
Db	2934	GCTTCAGTGCAGAACTCTCTGAACAAAGTAGAAGCTGAAGCGGCACCTGTTAACTGAGGAG	2993
Qy	2778	GTCACTCAAAAAGACATCATTCG---AAGAGAAACTCTGTGTCTCACCCCGACGCTTACCA	2834
Db	2994	GTATTGGAAGAGAGTAAATTGCAGAAAGAAACCCCCACCGTTATCTGAACCTCTCGCCA	3053
Qy	2835	GAGGGTAAAGATCCCATGACATGTGTCCACAGTGAAGTGGATTTCACTCTCAGAAAGCT	2894
Db	3054	GAGAAACAGAGAGGCCCGGGCGACCGGTGCTTTAGTAGGGCGGAATTGACCCCGAAGCT	3113
Qy	2895	GTGACGCCACAGAGACCTTCAGAGGCTCTCCGTACTGAAGAAAGTTACCGAAGCATCGGG	2954
Db	3114	GTGACAGCTCGAGAACTCGCAGGGCCATTGGGTTTCCGAAGAAAGAACCGAAGCATCTGCT	3173
Qy	2955	GCCGAAGAGACACAGACATGTGTCCGCAAGTTTCCCACTGACTGACTCCCCAGACACC	3014
Db	3174	GCTGAAGAGACCAAGAAATGGTGTCTCAGAGGTCTCCCAAGTTAACCGACTCCCCAGACACC	3233
Qy	3015	ACAGAGAAAGCCACCCAGTTACGAGGCTGAGAGTGGTGTCTAGATACAGAAAGAGAG	3074
Db	3234	ACAGAGAGGCCACTCCGTTGACGAGGTGGAAGGTGGGTACCTGACATAGAAGAGCAA	3293
Qy	3075	GAGCGCCAGAGCGAGGCCATCTCCAAAGCGTTGACAGCAAGGTGAAGAGGAGTCCCGAG	3134
Db	3294	GAGAGGCGGACTCAAGAGGTCTCCAGGCGAGTGGCAGAAAAAGTGAAGAGAGAAATCCCGAG	3353
Qy	3135	GTGCTCTCAACC---CAGACTGTGCAGAGAAACGGGGTCAAAAAGCAGCTGAGAAAGGTTGAG	3191
Db	3354	CTGCCTGGCACCGGTGGGCCAGAGATGTGCTTCAGCTGTGTCAGAGAGCAGAGGCGAGAA	3413
Qy	3192	GAGGTAGGAGGACTCCGAAGTGTGGCTTCCGAGAAAGAGAAAGGACGTTTATGTCGAGAA	3251
Db	3414	AGACCAAGAGAGCAGGCTGAAGCGCTCGGGTCTGAAGAAAGAGACGGAATGTAGTGTGAAA	3473
Qy	3252	GGACCCCTGCAGAGACTGGAGCTGAGCATCTTCGACAGGGCTCTGAGACTGGACAGGCT	3311
Db	3474	GTAGATCTCAGAGGGCAAAACTGAGCCTTTTACAAAGGGAAAGGTGGTGGGCGACACC	3533

QY	3312	ACTCCAGAGACCTTG---	AA GTTCTGAAGTCACGGCAGATGTAGAC-----CAT	3355
Db	3534	ACCCGAGAAAGCTTTTGAAA	AAAGCTCCTCAAGTCACAGAGAGCATAGAGTC	3593
QY	3360	GTCCACCACTGCGCA-----	GGTTATCAAGTCCACGACGTCATGGAA	3401
Db	3594	GTAAACCACTTGTCAAGCGNA	ACCTTAGCTGGGGTAAATCAACAGGAGATG	3533
QY	3402	CAGCGCGTGGCCCCCTGAGT	CAATCCGAAACCTTTGACAGACAGTGAACAAAT	3461
Db	3654	CAGGCTATCCCCCTGACTCG	TGGGAAACCCCTACAGACAGTGTAGACTGAT	3713
QY	3462	CCCTTAGCAGATTCAGACACT	GCAGATGGGACACAGCAAGATGAACCATTTG	3521
Db	3714	CCCGTAGCCGACTTTGACGCA	CAACCCAGAAAGACGAGATTTGTGGAAAT	3773
QY	3522	GACAGTAAAGCCACTGCAGCT	GTGCAGGAGTCAACAGGTTCACAGAAAGAGCG	3581
Db	3774	GAGGAGAAATGA---GGTGC	ATCTGGTACACAGTCAGGGGCAACAGACAGAG	3830
QY	3582	GCTCAGAAAGAGGAGCCTTC	GCACACTACCTAAATGTTTCCAGCCACAGAA	3641
Db	3831	GCA CAGAAAGAGAGGCTCC	CAGCACCTCCAGTTTGTGTTCCAGGAAGAA	3890
QY	3642	GAAGAACCCAGAAAGATGTT	CTTGAACCTACACAGCAAGAGCTTACTGCTG	3701
Db	3891	CAATCAAAAGATGGAAGAC	ACTCTAGAGCATACAGATAAAGAGGTGTCA	3950
QY	3702	CCCGTCTCGCAAGACTGAGGT	TGGGTCAAGAGGGTGA-----GGTTGACTG	3752
Db	3951	TCCATCTGTCAAAAGACTG	AGGGGACTCAAGAGGCTGACAGTATGCTGAT	4010
QY	3753	GATGGAGAAAAGTCAAGAA	GAAACACAGGAGGTGTTGTGTACACTCTGG-----	3806
Db	4011	AAAGCACTACCAATTTTCG	NAAGCACTTGAGGGCTCTATAGACACAGGC	4070
QY	3807	AGTCAAAAGGCTGCTGATGT	GCATATGACGTGAAGTGTGGAGTGGCCGGTGT	3866
Db	4071	CGGAAAAGGTCACTGAAGT	TGCGCTTAAAGGTGAAGGCACAGAAAGCTGA	4130
QY	3867	GAAAGCAGAGTACTGAAGT	GCAGAG-----TCTTAGCTCGAGAGGGA	3911
Db	4131	AAGGATGATGCTTTGAAC	TGCAGAGTCAAGTATGCTCTCCATCCCGCTG	4190
QY	3912	GAGATGAAAACCTCAGCTT	GAAAAGGAGAAAAGGGAGACAAAGCCACAG	3971
Db	4191	GAGATGTAGTTCAGTTCG	NAAGGAGAAAACAGAAAGCAGAGCCACCCAT	4250
QY	3972	GAAAG---TGAGCAGGAAA	CAGCCGCTCTGTAGCATGAAGGAACCTACG	4028
Db	4251	GAGAAAGCTTGAGCAGAAA	CAGCTGTTACCGTATCTGAAGAGGTCAAGT	4310
QY	4029	CTGACACTTCACATGCCAG	CTCAGAGAGGGGAGGCACTGGGNAAGCCTT	4088
Db	4311	CAGACGTGAATGTGCCAT	CAATAGATGGGCAAAAGGAAGTCAACAGT	4370
QY	4089	CTTTCTC---TCCCAGAC	CAAGCAAGCAGGTTCATAGAGTTTCAAGTT	4145
Db	4371	CCTCTCTCCCTAGTCAAG	GAGGAGGCAGTATGCACCAAAATTCAGTT	4430
QY	4146	GACACAAACAGTCACTCAA	ACAGCAAGCTGTGGAAAA-----GGTCAT	4199
Db	4431	GAGGCATCAATTCACTTA	CACCGGCTGCAGAGGAGGAAAAGGTCTTT	4490
QY	4200	GTGATTTTCAGAGACAGGT	GAAAGTCCAGAGTGTGTAGGTGCACACTTA	4259
Db	4491	AACATTTTAGAAAACAGGT	GAAACGTTTGGAGCCCTCAGGTGCACATATT	4550
QY	4260	AAGTCTCTCTCAACGGGT	GGCACTGGACTCTTCAGCATGCAGAGGACA	4319
Db	4551	AAATCTCTGAAAANAAT	GAGACTTTTCCGCTCATCCAGGGGAAGATG	4610
QY	4320	GGGCCTGAGTCTCAGGC	AGAAATCCATCCCAATCATAGTAACTCTC	4379

Db 4611 GGGCCCGCTGTGAGGCAAAATCGACACCGATGATGATGCTCTACTTCAACGAAAGGC 4670
 Qy 4380 CTACATCTCTGACTTCAAGAGAAATTAAGCGCATCCAGAGAGCGATCAGAGAGAG 4439
 Db 4671 TTAAGTTCGACCTGGAAAGAGAGAAACCATCATCTGAGAGTGAAGTCAAGTAGTC 4730
 Qy 4440 GACAAAGCAGATGCTGTGCTGATGCTGACGCAAGAGAGATGACGCAATCGAAAAATGC 4499
 Db 4731 GATGAGCAGGTGTCTTGGC-----AGGAGGTCAAGATGATGATGCAATTG---AGAT 4781
 Qy 4500 CTCAAGCTGAACTGAGATCTGGAATCTGAGATTAAGCAAAAGATTTGCTGAC 4559
 Db 4782 TTAGAGCTGAAATGGATTTTGGATTTTGGATCTGACCAAAAGATTAATTTTCCAAAC 4841
 Qy 4560 GTCAATTCAGACAGCGCTTGACAGTTGCACTGACCTAC---AGAAACAGCCCGCAAACTCAT 4616
 Db 4842 ATCATCCAGACAGCGCTTGACAGTTTGTAGTACAGAAAGAAACGCAACGAAATGTTG 4901
 Qy 4617 GCTTATGATTCAGACAGCCAGGTTCT-----CTGCATGACGCTTGAACG 4660
 Db 4902 ACGTCTGATTTACAGACACAGCTCAGATTAAGCTGACAGCCAGAGACGCTGACAG 4961
 Qy 4661 CAGGAGCGCCCAACAGATGCTGCAAAAATGAAGATGCCAAGATGAACACCCAGTGC 4720
 Db 4962 GAAACGAGAAAGAGAGAGAGAACTCAGGCTCTGACAGAGATGAACACCAATTAAT 5021
 Qy 4721 GCAGCCCA-----GAGAGACTTGAAGTCTGACCCGTTCTGA 4759
 Db 5022 TCAGCCAAAGAGAGATCAGAGTCAACCGCAGTGGGCAAGCATTCTGATTTTCCAA 5081
 Qy 4760 GGCATGAGCTCAGCTCGGAAATGCTTGCCTGCTTGCAGTTGAAGCCGCGTCA 4818
 Db 5082 GACATGAGTGAAGCTTCAGAAAGAACCATGACTTTGAAGTGAAGGTTCACTGTAAT 5141
 Qy 4819 ----- 4818
 Db 5142 GATCAGCAGCTGGAAGAGTCTCTCCCATCTGAGAAAGAGAGGTGAGCTGGAACA 5201
 Qy 4819 -----AGTAAGCATTTAGAGAGTGCCT 4840
 Db 5202 AAGTCTGTGCCAGAAATGATGTCATGCTTGTAGCAAGAAAGATTAAGAAATGACATA 5261
 Qy 4841 CCTCAACCCCAAGA--TCCAAAGAGACATGCTGATGAGCTCAGCTCCAAAGCTTA 4898
 Db 5262 GTTGAACCGAAAGAGATGAATAAAGTGATGATGATGATGATGATGATGATGATGAT 5321
 Qy 4899 GCGCAGCAGAGAGCCAGTGCCTCTGAAACCTAACCAAGATCCCGACAGACCAACGGA 4958
 Db 5322 GCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5381
 Qy 4959 CCAAGCTTAACCGAGAGAGGCGCATCCCGCAAAAGTTGAGTCCAGAGAAAGAAATGCTT 5018
 Db 5382 CCAAAACCAAAAGAGAGAGATGCGCCAGAAAGTGAATTTGACAGAAAGAAAGATGCGC 5441
 Qy 5019 ACCAAGTCAATCAAG 5078
 Db 5442 AGTGAATCAGATTAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 5501
 Qy 5079 GACCTGTGAGATCTCAAGATGTTAGTTG-----CTCATTTGATC 5117
 Db 5502 GAATGTGCAAGTCAAGAACTTACAGATTTTAAACATCATGCAAGTTAACTCATTTGCT 5561
 Qy 5118 ATCTGTAAGACAGAGATGTAAGAAACAGTCAAGAAACA-----GATGCTGTGTTGGAG 5173
 Db 5562 GTTTGGAAGACCAAGATGTAAGAGACAGATGTAAGAAAGAAAGATGATGATGATGATGAT 5620
 Qy 5174 CTTGAGACCAAGATTTCAAGAGCCATGAGATCCAGAGACAGAGCCGTCATGATTTCC 5233
 Db 5621 CTGAAGACCAAGATTTCAAG-ACCTTGAAGATTTGAGAGACAGGACATCACTGATCTCA 5679
 Qy 5234 ACCCAGTGAAGACCCCGCAATTTCTGAGGCTTCACTGGGAGCTGAGGCACTTAACATT 5293

Db 5680 TTTCAGAGAGC-CCCTGACATCTGAGGCTTCAATCAGAGAGTGAAGCATTTAATTT 5738
 Qy 5294 TCTCTGTTTCAAGCTGCTTTGATTTGCCCCCTTGATGCGCTCGTGTATTTCTAATTT 5353
 Db 5739 TCTCTTTCAGAGACCAACCTAATTTTCTGATTAAC----- 5778
 Qy 5354 TCTCTGTTTCAAGAGCTGCTTTGATTTTGCCCTTGATGCGCTGCTGATTTTGGATTTA 5413
 Db 5779 -----CATATAATTTGATTTA 5796
 Qy 5414 AGTCTGCGTCTTCAACCTGGAAC-CAATTTGCAATCTGATCTGATCTCAATTTCAACT 5472
 Db 5797 AGTCTCAATTTCTTAACCTGGAACCTGAGTGGCAATCTGATCTGATCTGGAACCT 5856
 Qy 5473 GGAAGATCTCTTTATGATTTATATGATGTTTATATGATGTTTATGATGTTTATGATGTT 5532
 Db 5857 GGAATATCATTTCTTAATATTTATATGATGTTTATGATGTTTATGATGTTTATGATGTT 5913
 Qy 5533 TGTATATTTTCTTAAGTT-----AAGCAGATCTTTTGTATTA 5575
 Db 5914 TGTATATTTTCTTAATGTTTAAAGAAATGTCAGATATCTAATGCTTTTGTATTA 5973
 Qy 5576 TGCATATATTAACGGTGTGACCATAGCAGACGCTTTGAAAGCTCCAAAGCTCAACTG 5635
 Db 5974 CACAGTATATGATGGGAGATGTCATATGTCAGAGCTTTGGGAGCTTTAAGCTCAAGTTA 6033
 Qy 5636 TTAACCTGAGCAACAGATTAATTC-----CTGGCAAGAGAGACAGTCTTTT 5686
 Db 6034 TATTAACCAAAAGAGAGAGCTCTAGATGTAACATTCCTGATCAAGATCAATTTCT 6093
 Qy 5687 TTAAGTTTACGATGCTTAATGATGTCGGGCTTCTAGCTCTGAAAGGTTGTTTCTCT 5746
 Db 6094 TTAATTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6153
 Qy 5747 ATGCAAGCGAGCTCAGAAATTAATAAACCCTATTTGAAACATCCAGATGTCCTCAATTT 5806
 Db 6154 ATTACAGGAGATGTCAGAAATTAATAAAGCTTTGAAACATTAAGATGTTCTATTTGTC 6213
 Qy 5807 ACCATGATTTTTCCTCTTTTCTTATCTATCTGATGATGATGATGATGATGATGATGATGAT 5866
 Db 6214 ATTGGAAATTTTGC-----TTCTTAACCAAGTGGGATTAAGAAAGATTAATTTCT 6266
 Qy 5867 TGTGAGA-----TTAAGCCCTGCTTAAATGATGACAAATGAT 5909
 Db 6267 GGTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6326
 Qy 5910 GTGCTTAAGCCATGAGA-----TGTTCCTAATGACAGAGATCTGT 5953
 Db 6327 GTGCTTAATCTGAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6386
 Qy 5954 TGTAGTTTTTGTATGATCTTCTATGCTGAGACCGAATTCATATGACAGATGGAAGTG 6013
 Db 6387 TGTAAAG--TTTTGATTTCTACTTATATGCTGAGACGTCATTCACATGSCATGAATA 6444
 Qy 6014 AGTCTGTTCTTTACAGATGATGATTTTGAATGATGATGATGATGATGATGATGATGATGAT 6073
 Db 6445 AGTCAAGTCTTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6504
 Qy 6074 TGCCCTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6133
 Db 6505 TGCCATTCCTTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6564
 Qy 6134 CTGATTTAAATTA 6147
 Db 6565 CTGATTTAAATTA 6578
 RESULT 8
 AAS85205
 ID AAS85205 standard; cDNA; 6886 BP.
 XX
 AC AAS85205;

DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #21009.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI P-PSDB; ABG21018.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX Claim 1; SEQ ID No 21009; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 6886 BP; 2142 A; 1483 C; 1908 G; 1352 T; 1 other;
Query Match 31.2%; Score 1920.6; DB 23; Length 6886;
Best Local Similarity 67.3%; Pred. No. 0;
Matches 3255; Conservative 0; Mismatches 1334; Indels 244; Gaps 29;
QY 3 GGCGCGGAGTAGAAGACCCACTGAGCCATGGCGCGAGGAGTTCACCGAGACGCG 62
DB 162 GGTGAGCGCGGAGGAGTTCGGAGGAGGAGTTCAGCGCGGAGTTCACCGAGACGCG 221
QY 63 AGCCCGAGCAGCGG---GCGGGGAGCGACACCGAGCGAGTGGTCTCAGTGGCCAT 119
DB 222 AGCCCGAGCAGCGCGCCGAGGAGGAGTTCACCGCGGCTGAGCCCGAGCGCGCGCGC 281
QY 120 GGGCCCGCAGCTGAAGCCTCGGGA-----GCAGCTGGAGACCCCGCGAGCGCGAC 170
DB 282 GGGCCCTCGCGGAGGCGGCGCCAGACACCGCGGACCCCGCCATCGCTGCTCGGAC 341
QY 171 CCGGCCCAACGAAGTCCCAAGAGAAATGGCCAGGCTGTCTTCTGTCAACGCGGTAGCTGAA 230

DB 342 CCGGCCACCAAGCTCCTACAGAAAGAAATGGTCAGCTGTCCACCATCATGCGGTAGCTGAG 401
QY 231 CAAGGAGATGCTCCATGTCCAAAGAGGAAACCCAGGAGG----- 268
DB 402 CAAGATGAGCTCAGCCCTCCAGGAGGGGTGACCTAAATGGCCAGAAAGGAGCCCTGAACGGT 461
QY 269 -----GCAGGAGGAAAGAGTCTGTGATGAGGATGTTGGACACGGA 308
DB 462 CAAGGAGCCCTAAACAGCCAGGAGGAGAAAGAGTCAATGTCCGGCAGGTTGATCACATA 521
QY 309 GAGTCAGAGATGTGAGAGAAAAGACCCGAGTTCAGAAATGGCGGCAACTCCACAGCT 369
DB 522 GACTCTGAAGATGTGAGCGACAGAGACTCCGATACAGAGATGACTACTAAGTCAGCGGTA 581
QY 369 GTTGAAGATATCAAAAGGATGGCAGGAGGAGACATCAGAAATAAATTTGAACAGATCCCT 428
DB 582 GTTCAAGATATCAGATGATGGCAGGAGGAGACACCCGATATATTCGAAACAGATTCTCT 641
QY 429 GCTTCAGAAAACAATGTGGAAGAAATGGTACAGCCTGCTGAGTCCCAGGCTAATGATGTT 488
DB 642 TCTTCAGAAAGCAATTTAGAAGAGCTAAACACACCCACTCAGTCCCAGGCTAATGATATT 701
QY 489 GCGTTCAAGAAAGTATTAAATTTGTTTAAATTTACGGTGAAGAGGATAAAAAT 548
DB 702 GGAATTTAAGAAAGGTGTTTAAAGTTTGTGGCTTTAAATTTCACTGTGAAAAGGATAAGACA 761
QY 549 GAAAAGTCAGATACTGTCCAACTACTCAGTCTCAAGAGGATGAAGCGCAAGGGCGAGAA 608
DB 762 GAGAGCCTGACACTGTCCAGCTACTCAGTGTGAAGAAAGATGAAGGGGAGGAGCAGCA 821
QY 609 GCCTGTCTCGAGCTGGAGACCAACAGGAGCCAGTGTGGAGACTGCGCTCGGAGAGTCA 668
DB 822 -----GGGCTCGCGACACACAGGACCCAGCCTTGGGCT-----GGAGAGCA 866
QY 669 GATTCAAAGAAAGTGAAGTGAAGCAATCCACAGAGAGCAAGAGCACCTTGAAGCAA 728
DB 867 GCATCAAAAGAAAGCGAAACCAACCAATCTACAGAGAAACCCGAGAGACCTTGAAGCGT 926
QY 729 GAACAGAGCAGCAGAGAA--TCCCTCTTCAAGCCGAACTCTGATCAAGCGCTCAGGAGGA 787
DB 927 GAGCAAGCCACCGCAAAATTTCTCCCAAGCGGAATCTGGCCAA---CGATGGAGGA 983
QY 788 AGCCAAGATGAAGGAGAGAGAAACAAAGAGAAAGAGCCCAAGTCCCAGAAATCCCC 847
DB 984 ATCAAGAGAGGAGGAGAGAGAGAAACAGAAAGAAAGACCTAGCAAGTCTCAGAAATCTCC 1043
QY 848 GAGCAGCCAGTCAACAGTGAAGACATCTTCTTCAAGAAAGTCTTCTCACTCAGCGTTG 907
DB 1044 GACTAGTCCCGTGACCCAGTGAACAGGATCAACCTTTCAAAAAATTTCTTCACTCAAGTTG 1103
QY 908 GGCCTGCTGGCGCAAGAGACCAAGTTCAGAAATCAAAAGAGGATGATCTCGAAACTGC 967
DB 1104 GGCCTGCTGGCGCAAAAGACCAAGTTCAGGAAAGCCGAGGAGGATGAAGTGAAGTTT 1163
QY 968 CGAGAGAGAAAGGAGCAGAGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1027
DB 1164 AG 1223
QY 1028 GCAGGCTCGAGGA-----GCAGGAGCCGCGAGAGGA 1060
DB 1224 GGTTCGCTCCGAGAAACTGACCCCTCCGAGCAAGCCCAAGCCCAAGGCGCGAGAGAG 1283
QY 1061 CACAGACCCAGGCGAGTGTTCAGCAGACTACAGAGAGGTTGAGCTGCTTTCGAAAGCCA 1120
DB 1284 TGCCCAAGAGCCCGGTTTATCAGCTGAATATGAGAAAGTTGAGCTGCTTTCAGAGGAGCA 1343
QY 1121 GGTTCGTGACCTGAGGAGCATCGTCAGAGAGAGTGTGCTCTCTTTCGCAACGGAAGTGT 1180
DB 1344 AGTCAGTGGCTCGCAGGAGCTTCTGAGAGAGAAACCTGCTCCGTTGGCGACAGAGTGT 1403
QY 1181 TGATGAGAGATGGAAAGCCCAAGAA---GTTTGTGAGAGGTCACAGTGAAGCACCGT 1237

Qy	3358	ATGTGCCACGTGCCA-----GGTTATCAAGCTTCAGCAGCTGTATGG	3399
Db	3597	TTGTTAAACCATTTGTCGAAGCGAAACCTTAGCTGGGGTAAAAATCA	3656
Qy	3400	AACAGGCGGTGGCCCTGAGTCAATCGAAACCTTGACAGACAGTGA	3459
Db	3657	AACAGGCTATCCCCCTTGACTCGGTGGAAACCCCTACAGACAGTG	3716
Qy	3460	CTCCCTTAGCAGATTACAGACACTGCGAGATGGGACACAGCAAGAT	3519
Db	3717	CCCCGTTAGCCGACTTTGACGCACACAGGCACAACCCAGAAAGAC	3776
Qy	3520	AGGACAGTAAAGCCACTGACAGCTGTGACGCGATGTCACAGGTCA	3579
Db	3777	ATGAGGAGAAATGAGGTGCGATCTGCTGTACCGACTCAGGGGGGCA	3836
Qy	3580	CTGCTCAGAAGAGGAGCCCTTGACACTACCTTAATATGTTTCCAGC	3639
Db	3837	CTGCACAGAAGAGAGAGCCCTCCAGACCTTCAGATTTTGTGTTCC	3896
Qy	3640	GGGAAGAACCCAGGAAGAGATGTTCTTTGAACCTACACAGCAAGAG	3699
Db	3897	AACAAATCAAGATATGGAAAGACACTCTAGAGCATACAGATAAAG	3956
Qy	3700	TGCCCGTTCTGGCAAGACTGAGGTGGGTCAAGAGGGTGA-----	3750
Db	3957	TATCCATTCTGTCAAAGACTGAGGGGACTCAAGAGGCTGACCCAGT	4016
Qy	3751	TGGATGGAGAAAAGTCAAGAAGACACGAGAGGTGTTGTACACTCTG	3804
Db	4017	CCAAAGACGTACCATTTTTCGAAGGACTTGGGGGGTCTATAGACA	4076
Qy	3805	ACAGTCAAAAAGGCTGCTGATGTGCATATGACAGTGAAGTGAATG	3864
Db	4077	GTCCGGAAAAGGTCACATGAAGTTGCCCTTAAAGGTGAAGGACAG	4136
Qy	3865	AGGAAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCTCGG	3909
Db	4137	AAAAGGATGATGCTCTTGAACCTGCAGAGTCACGCTAAGTCTCTC	4196
Qy	3910	GAGAGATGGAAACTGACGTTTGAAGGAGAAA---AGGAGACAAAG	3966
Db	4197	GAGAGATGGTAGTTTCAAGTTCGAAGGGAGGAAATCCAGTAGCAG	4256
Qy	3967	GTGAAG---AAGGTGACGAGGAAAACAGCGCTCCTTGAGCTGAAG	4023
Db	4257	ATGNAAGAAAGCTTGAACAGCAAAACAGCTGTACCGTATCTGAAG	4316
Qy	4024	CAGTCTTGACACTTGAATGCCCGAGCTCAGAGAGGGGGAAGGCACT	4083
Db	4317	TCCTCCAGACAGTGAATGTGCCCATCATAGATGGGGCAAGGAAGT	4376
Qy	4084	GAAGCCCTTCTC---TCCCNAGCAAGACAAAGCAGGTTGCATGAGG	4140
Db	4377	GAAGCCCTCTCCCTCGCTAGSTCAAGAGGGAGGCAATGACACCAAA	4436
Qy	4141	GCCTGGACACAAACAGTCACTCAAACAGCAGAAAGCTGTGMAAAGT	4194
Db	4437	GCTCTGAGGCATCATTTCACTCTTACAGCGGCTGCAGAGAGGAAGG	4496
Qy	4195	CGGTTGTGATTTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTG	4254
Db	4497	CTGCCAACATTTTATAGAAAACAGGTGAAACGTTGGAGCCTGCAGG	4556
Qy	4255	CTGAGAAGTCTCTGCAACGGGTGGCCACTGACACTCTTCAGCATGC	4314
Db	4557	AAGAGAAATCCTCTGAAAAAATAAGAAAGCTTTTCGCGCTCATCC	4616
Qy	4315	CCCTGGGGCTGAGTCTCAGGCAGAAATCCATCCCAATCATATAGTA	4374
Db	4617	CCACAGGGCCGAGCTGTACGGCAAAATCGACACCACTGATGATGTC	4676
Qy	4375	GCACCCCTACATCTCTGACCTACAAGGAGAAATTAAGCGCAATCC	4434

Db	4677	AAGGCTTAAGTTC	CGACCTCGAAGGAGAGAAACCACTACTGAAGTGGAAAGTCAGATG	4736
Qy	4435	AAGAGGACAAGCCAGATGCTGGTCCCTGCTGACGGCAAGCAGAGTACAGCAATCGGAA	4494	
Db	4737	AAGTCGATGAGCAGGTGCTTGCC-----AGAGGTCAAAGTGTAGTCAATTTGAGG	4790	
Qy	4495	AAGTCCTCAAGGGCTGAACCTTGAGATCCTGGAACTTGAGAGTAAAGCAACAAGA-TTGTG	4553	
Db	4791	AGGATTTAGAGCCTGAAATGGGATTTGGAACTTGAGACCAAAAGCAGTAAAACTTGTC	4850	
Qy	4554	CTGAACGTCATTCAGACAGCGGTGACCAAG-TTCGCACTGTAC---AGAAACAGCCCCCGA	4609	
Db	4851	CAAAACATCATCCAGACAGCGGTGACCAAGTTTGTACGTACAGAAAGAAACAGCCACCGA	4910	
Qy	4610	AACTCATGCTTATGATTCACAGACCCAGGTTCC	4642	
Db	4911	AATGTTGAGCTGTAGTTACAGACACAAGCTTC	4943	
RESULT 9				
ABK63316/c				
ID	ABK63316 standard; cDNA; 563 BP.			
XX	AC			
XX	AC			
XX	ABK63316;			
DT	18-JUN-2002 (first entry)			
XX	Rat sequence differentially expressed in response to a hepatotoxin #1223			
DE	Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;			
KW	differential expression; centrilobular necrosis; steatosis.			
XX	OS			
XX	Rattus norvegicus.			
XX	WO200210453-A2.			
FN	07-FEB-2002.			
PD	30-JUL-2001; 2001WO-US23872.			
XX	31-JUL-2000; 2000US-222040P.			
PR	02-NOV-2000; 2000US-244880P.			
PR	11-MAY-2001; 2001US-290029P.			
PR	15-MAY-2001; 2001US-290645P.			
PR	22-MAY-2001; 2001US-292336P.			
PR	06-JUN-2001; 2001US-295798P.			
PR	13-JUN-2001; 2001US-297457P.			
PR	19-JUN-2001; 2001US-298884P.			
PR	09-JUL-2001; 2001US-303459P.			
XX	(GENE-) GENE LOGIC INC.			
PA	Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;			
XX	WPI; 2002-241625/29.			
XX	Predicting toxic effects of compounds or the progression of these toxic			
PT	effects by determining the changes in gene expression in tissues or			
PT	cells exposed to the toxin and comparing these to gene expression in			
PT	unexposed tissues or cells -			
XX	Claim 1; Seq ID No 1223; 239pp; English.			
PS	The invention relates to methods for predicting toxic effects of			
CC	compounds or the progression of these toxic effects by determining the			
CC	global changes in gene expression in tissues or cells exposed to the			
CC	toxin and comparing these to gene expression in unexposed tissues or			
CC	cells. Also included are methods of predicting at least one toxic			
CC	effect of a compound or progression of a toxic effect, preferably the			
CC	hepatotoxicity of a compound, comprising detecting the level of			
CC	expression in a tissue or cell sample exposed to the compound of two or			
CC	more genes listed in the specification. where differential expression of			

CC the genes is indicative of at least one toxic effect or progression.
 CC The method can also be used to identify an agent which modulates the
 CC toxic response and predict cellular pathways that a compound modulates
 CC in a cell. The methods utilise a set of at least two probes (on a solid
 CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridises to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity
 CC is characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent.
 CC
 XX

Sequence 563 BP; 191 A; 107 C; 110 G; 155 T; 0 other;

Query Match 8.9%; Score 550; DB 24; Length 563;
 Best Local Similarity 99.1%; Pred. No. 8e-132;
 Matches 553; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

5601 ATAGGAGCGCTTTGAAAAGCTCCAGCCCACTGAACCTGACCAAAACAGATACATT 5660
 563 AAAGGAGCGCTTTGAAAAGCTCCAGCCCACTGAACCTGACCAAAACAGATACATT 504
 5661 CCTGCAAGAGAGACAAGCTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 5720
 503 CCTGCAAGAGAGACAAGCTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 444
 5721 AGTCTCTGAAAAGGTGTTTCTTATGACAGAGAGCTGAGAAATATAAAACCCCATTT 5780
 443 AGTCTCTGAAAAGGTGTTTCTTATGACAGAGAGCTGAGAAATATAAAACCCCATTT 384
 5781 TGAACATCCAGAGATGCTCCCAATATTAACATATTTTCCCTTTTGGTAAATCCAG 5840
 383 TGAACATCCAGAGATGCTCCCAATATTAACATATTTTCCCTTTTGGTAAATCCAG 324
 5841 TCCAGGTTGAAAAGAGTCTCTCTGTGTGATTAAGCCCTGTCTTAAATGATATGA 5900
 323 TCCAGGTTGAAAAGAGTCTCTCTGTGTGATTAAGCCCTGTCTTAAATGATATGA 264
 5901 CAAATGAGTGTGCTTAAGGCCATGAGATGTTTCTTAATGACAAAGAAATCTGTGACT 5960
 263 CAAATGAGTGTGCTTAAGGCCATGAGATGTTTCTTAATGACAAAGAAATCTGTGACT 204
 5961 TTTTGTGATGACTCTTCTATGCTGAGACCAATCAATGAGATGAGAGTCTG 6020
 203 TTTTGTGATGACTCTTCTATGCTGAGACCAATCAATGAGATGAGAGTCTG 144
 6021 TTTTGTGATGACTCTTCTATGCTGAGATGAGAGTCTGAGATGAGAGTCTG 6080
 143 TTTTGTGATGACTCTTCTATGCTGAGATGAGAGTCTGAGATGAGAGTCTG 84
 6081 TTTTGTGATGACTCTTCTATGCTGAGATGAGAGTCTGAGATGAGAGTCTG 6140
 83 TTTTGTGATGACTCTTCTATGCTGAGATGAGAGTCTGAGATGAGAGTCTG 24
 6141 AAATAACATATTTGACT 6158
 23 AAATAACATATTTGACT 6

RESULT 10
 AAX40328
 ID AAX40328 standard; cDNA; 377 BP.
 XX

AAX40328;

16-JUN-1999 (first entry)

Human secreted protein 5' EST SEQ ID NO:115.

Human; secreted protein; EST; expressed sequence tag; diagnosis;
 forensic; gene therapy; chromosome mapping; signal peptide;
 upstream regulatory sequence; cytokine activity; cell proliferation;
 differentiation; haematopoiesis regulation; tissue growth regulation;
 reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 thrombolytic; anti-inflammatory; tumour inhibition; ds.

Homo sapiens.

WO9906439-A2.

11-FEB-1999.

31-JUL-1998; 98MO-IB01233.

01-AUG-1997; 97US-0904468.

(GIST) GENSFT.

Duclert A, Dumas Milne Edwards J, Lacroix B;

WPI; 1999-153700/13.

P-PSDB; AAY11610.

New nucleic acids encoding human secreted proteins - obtained from
 cDNA libraries derived from liver, lung, large intestine, colon,
 thyroid and pancreas tissue

Claim 1; Page 228-229; 398bp; English.

AAX40251 to AAX40397 represent 5' expressed sequence tags (ESTs) for
 human secreted proteins, and encode the proteins given in AAY11533 to
 AAY11679, respectively. The proteins given represent the signal peptide
 and an N-terminal fragment of a secreted protein. The nucleic acid
 sequences can be used for producing secreted human gene products. They
 can also be used to develop products for diagnosis and therapy. The
 proteins obtained may have cytokine activity, cell
 proliferation/differentiation activity, haematopoiesis regulating
 activity, tissue growth regulating activity, reproductive hormone
 regulating activity, chemotactic/chemokinetic activity, haemostatic and
 thrombolytic activity, receptor/ligand activity, anti-inflammatory
 activity, tumour inhibition activity or other activities. The products
 can be used in forensic, gene therapy and chromosome mapping procedures.
 The sequences can also be used for obtaining corresponding promoter
 sequences. The nucleic acids encoding the signal peptide can be used for
 directing extracellular secretion of a polypeptide or the insertion of a
 polypeptide into a membrane, or importing a polypeptide into a cell.

Sequence 377 BP; 111 A; 85 C; 113 G; 66 T; 2 other;

Query Match 2.8%; Score 172.2; DB 20; Length 377;
 Best Local Similarity 68.9%; Pred. No. 4.3e-34;
 Matches 253; Conservative 1; Mismatches 104; Indels 9; Gaps 1;

2455 AAGCCAGTGTGAAGACCTCAGGCGCAGTGAGATTAATGAGAGAGACCTATGTCACG 2514
 1 AAGCCAGTGTGAAGACCTCAGGCGCAGTGAGATTAATGAGAGAGACCTATGTCACG 60
 2515 CGGTGCTCTCTGTGATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 2565
 61 CGGTGCTCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 2566 CCCAGGGGAATACGAGAGTGCCTCCAGCTGCTGGGGCTGTGTACGTGTCCAGAGAGCTCA 2625
 121 CCCAATAAAGCGAGAGAGCGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 2626 GTAAGACTCTGTGTCCACACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2685

Db 181 GCGAGAGTCAGTTCATATGATGCGACAGCTGCTGCTGACGGACGAGGCGACTACCA 240
 Qy 2886 GTGTCGAAGACGGTCTCTTCCTGCGATATCCGTTCCGTTACACAGAACCTCTTGAACACA 2745
 Db 241 TTATTGAAGAAGGCTCTCTTCCTGCGATATCTGCTTCAGTCACAGAACCTCTTGAACAAG 300
 Qy 2746 CAGCGGAGAGCCACCTGTTGAGAGTTCACCTGAAAAAGACATCTTCAGAGAAG 2805
 Db 301 TAGAAGCTGAAGCCGCACTGTTAACTGAGGAGGTATTGGAAGAAGAGTATTGCGAGAAG 360
 Qy 2806 AAACCTCC 2812
 Db 361 AAGAAC 367
 RESULT 11
 ABK63568
 ID ABK63568 standard; cDNA; 178 BP.
 AC ABK63568;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Rat sequence differentially expressed in response to a hepatotoxin #1475.
 XX
 DE Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
 KW differential expression; centrilobular necrosis; steatosis.
 KW
 OS Rattus norvegicus.
 XX
 XX WO200210453-A2.
 XX
 PD 07-FEB-2002.
 XX
 XX 30-JUL-2001; 2001WO-US23872.
 XX
 PR 31-JUL-2000; 2000US-222040P.
 PR 02-NOV-2000; 2000US-244880P.
 PR 11-MAY-2001; 2001US-230029P.
 PR 15-MAY-2001; 2001US-290645P.
 PR 22-MAY-2001; 2001US-292336P.
 PR 06-JUN-2001; 2001US-295798P.
 PR 13-JUN-2001; 2001US-237457P.
 PR 19-JUN-2001; 2001US-298884P.
 PR 09-JUL-2001; 2001US-303459P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
 XX
 XX WPI; 2002-241625/29.
 XX
 XX Predicting toxic effects of compounds or the progression of these toxic
 PT effects by determining the changes in gene expression in tissues or
 PT cells exposed to the toxin and comparing these to gene expression in
 PT unexposed tissues or cells -
 XX
 PS Claim 1; Seq ID No 1475; 239pp; English.
 XX
 CC The invention relates to methods for predicting toxic effects of
 CC compounds or the progression of these toxic effects by determining the
 CC global changes in gene expression in tissues or cells exposed to the
 CC toxin and comparing these to gene expression in unexposed tissues or
 CC cells. Also included are methods of predicting at least one toxic
 CC effect of a compound or progression of a toxic effect, preferably the
 CC hepatotoxicity of a compound, comprising detecting the level of
 CC expression in a tissue or cell sample exposed to the compound of two or
 CC more genes listed in the specification, where differential expression of
 CC the genes is indicative of at least one toxic effect or progression.
 CC The method can also be used to identify an agent which modulates the
 CC toxic response and predict cellular pathways that a compound modulates
 CC in a cell. The methods utilise a set of at least two probes (on a solid

CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridises to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity
 CC is characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent.
 XX
 SQ Sequence 178 BP; 41 A; 30 C; 33 G; 74 T; 0 other;
 Query Match 2.6%; Score 162.2; DB 24; Length 178;
 Best Local Similarity 97.8%; Pred. No. 1.1e-31;
 Matches 175; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 Qy 5962 TTTTGTGATTCTCTTCTATGCTGGACCGAATTCATATGCAGATCGAAGTCTCTGT 6021
 Db 1 TTTTGTGATTCTCTTCTATGCTGGACCGAATTCATATGCAGATCGAAGTCTCTGT 60
 Qy 6022 TCCTTACAGATGGTATTTTGTATAGATAGTCTGCTCTGTATATCTGTCCCTT 6081
 Db 61 TCCTTACAGATGGTATTTTGTATAGATAGTCTGCTCTGTATATCTGT-CCCTT 119
 Qy 6082 CTTTAAAGAACATGTTGCAATTTGCTTTGGATAAAATGTTGATTTGACAACTGATTT 6140
 Db 120 CTTTAAAGAACATGTTGCAATTTGCTTTGGATAAAATGTTGATTTGACAACTGATTT 178
 RESULT 12
 AAV86620
 ID AAV86620 standard; cDNA; 483 BP.
 XX
 AC AAV86620;
 XX
 DT 27-APR-1999 (first entry)
 XX
 DE EST clone BF146.
 XX
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO9845435-A2.
 XX
 PD 15-OCT-1998.
 XX
 XX 10-APR-1998; 98WO-US06954.
 XX
 XX 10-APR-1997; 97US-0835913.
 XX
 XX (GEM) GENETICS INST INC.
 XX
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie LA, Spaulding V, Treacy M;
 XX
 XX WPI; 1999-070076/06.
 XX
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries
 XX

PS Claim 1; Page 294; 633pp; English.

XX This sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC hematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cachectin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene
CC therapy.

XX Sequence 483 BP; 156 A; 103 C; 133 G; 90 T; 1 other;

Query Match 1.7%; Score 105.2; DB 20; Length 483;
Best Local Similarity 60.4%; Pred. No. 1.1e-16;
Matches 236; Conservative 0; Mismatches 134; Indels 21; Gaps 3;

OY 3811 AAAAGCTGCTGATGACATATGACATGAGTGGAGTGGCGGTGTCAGGAAA 3870
DB 44 AAAAGTCTGAGTGAAGTTCCTTAAAGTGAAGGACAGAGAGCTGATGTAAGG 103
OY 3871 AGGAGTCTGAGTGAAGTGAAGG-----TCTTACCTGAGAGGAGAGAGA 3915
DB 104 ATGATGCTCTTGAAGTCAAGATCAGTACGCTAAGTCTCTCCATCCCGGTGAGAGAGAGA 163
OY 3916 TGGAAATGACGTTGAAAAGGAAAAGGAGAGACAAAGCCAGAGAGTGAAGAGAG 3975
DB 164 TGGTGTGTTCAAGTCAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 223
OY 3976 G---TGACAGAGAAACAGCGGCTCTGAGCATGAGAGAACTACAGGAGAGAGCTCTGA 4032
DB 224 AGCTTGAAGACAAACAGTGTACCGAATCTGAAAGAGTCAAGTCAAGCTCTCTCAGA 283
OY 4033 CACTTGAAGTGGCCAGCTCAGAGAGGAGGAGCACTGGGAAGCTTGGAGAGAGCCCTT 4092
DB 284 CAGTGAATGTGCTCATATGATGGGGCAAGAGATCGAGAGTTTGAAGAGAGCCCTC 343
OY 4093 CTC---TCCAGAGCCAGAGCAAGAGGTTGCAATGAGAGTTCAAGTCAAGCTTGGACA 4149
DB 344 CTCCTGCTAGTCAAGAGAGAGAGAGTATGCAACCAATTCAGAGCTCTGAGAG 403
OY 4150 CAACAGTCACTCAAG 4180
DB 404 CATCATTCATCTTAACAGCGGCTGAGAGAGA 434

RESULT 13
ABV56830/c
ID ABV56830 standard; cDNA; 459-BP.

XX ABV56830;

XX 17-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 56821.

XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.

XX

PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE,
PI WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer

XX Claim 1; Page 10950; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 459 BP; 143 A; 99 C; 59 G; 158 T; 0 other;

Query Match 1.7%; Score 101.8; DB 23; Length 459;
Best Local Similarity 70.5%; Pred. No. 8.3e-16;
Matches 136; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

OY 5963 TTTTGAATGTACTCTTCTATGCTGACCGAATTCATATGACATGCAAGTGAAGTCTGTT 6022
DB 235 TTTTGAATTTTACTTTTAAAGTGGCGGCAATTTACACAGAGATGAATAAGTCAGTT 176
OY 6023 CTTTACAGATGATTTTGTATGATGATGAGTGTGCTGTATATCTGTGCCCTTC 6082
DB 175 TTTTACAAAGGAGATTTGAAGAAAACTGAGATGGGTTGGCCAAATTGGGCTTTT 116
OY 6083 TTTTGAAGCAATGTTGATTTGTTCTTGGATTAATTTGATTTGACATGATTTAA 6142
DB 115 TTTTGAAGCAAAAGTGGCCACCATTCATTTGATTAAGTTGATTTGACATGATTTAA 56
OY 6143 ATTAACATATTGG 6155
DB 55 AAAAAAAATTGG 43

RESULT 14
AAI86887
ID AAI86887 standard; cDNA; 400-BP.

XX AAI86887;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 6947.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukemia;
KW nervous system disorders; arthritis; inflammation; ss.

XX


```
OS Homo sapiens.
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 19-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
DR P-PSDB; AAO06956.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 1; SEQ ID NO 6947; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 400 BP; 138 A; 80 C; 90 G; 92 T; 0 other;

Query Match 1.5%; Score 95.4; DB 22; Length 400;
Best Local Similarity 63.3%; Pred. No. 3.5e-14;
Matches 241; Conservative 0; Mismatches 111; Indels 29; Gaps 5;

QY 4976 GGGCGATCCCCAAAAGTTCAGGTCCAGGAAGAGAAATGCTTACCAAGTCAGTCAAGA 5035
DB 1 GGAGGATGCCCGAGGAGTAGAATTGCAGGAGGAGGAAAGTGCACAGTGAATCAGATAAGC 60
QY 5036 GAACAAGGCCGAGCAGAGAGACCTGCAGGAGCCAAAGGAGACCTGGCAGAAATCCTA 5095
DB 61 GATCACACCCCACTCAGAGGAGGATTTTCAGAAACAGAGAGAGAAATCTGCCGAGTCTG 120
QY 5096 AGATGTTAGTTG-----CTCATTTGATATCTGTAAGACCGAA 5133
DB 121 AACTTACAGAAATCTTAAACATGCTGCAGTTAAACTCATTTGGTGTGGAAGACCGAA 180
QY 5134 TGTGMAAACAAAGTCACAGAACAG-ATGCTGCTCTGTGGACCTTGAGACCAAGATTT 5189
DB 181 TGTGAGACCCGGGATAGAGAAATGAATGCTGCTCTGATA-CTGAAGACCAACAATTT 239
QY 5190 CAGAGCCCATGAGATCCAGAGAGAGAGGCGGCTCCAAATGATTTCCACCAGTAGAGCACCC 5249
DB 240 CAGAACCT-TGAGAAATGGAGAGCAGGCACATCAATGATATCATTTCTAGAGC-CCC 297
QY 5250 CGACAATTCAGGCTTCATCGGAGCTAGAGCCAGGCTAACATTTCTCTGTTTCAAGACT 5309
DB 298 TGACGATCTCGAGGCTTCATCAGGAGCTAGAGCCATTTAAACATTTCTCTTTTCAAGACC 357
QY 5310 GCCTTTGATTTGCCCTTGAT 5330
DB 358 AACCTACAAATTTTCTCTTGAT 378

RESULT 15
AAV86330
ID AAV86330 standard; cDNA; 339 BP.
XX
XX AAV86330;
AC
XX 27-APR-1999 (first entry)
DT
XX EST clone AJ169.
DE
XX
XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
XX Homo sapiens.
XX
XX WO9845435-A2.
PN
XX 15-OCT-1998.
PD
XX
XX 10-APR-1998; 98WO-US06954.
PF
XX
XX 10-APR-1997; 97US-0835913.
PR
XX (GEMY ) GENETICS INST INC.
PA
XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI Racie LA, Spaulding V, Treacy M;
XX WPI; 1999-070076/06.
DR
XX New polynucleotides encoding human secreted proteins - derived from
XX e.g. human blood, kidney, foetal lung, placenta, testes, brain,
XX ovary, pituitary, retina and colon cDNA libraries
XX
XX Claim 1; Page 198; 633pp; English.
XX
CC This sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene
CC therapy.
XX
SQ Sequence 339 BP; 109 A; 74 C; 91 G; 65 T; 0 other;

Query Match 1.5%; Score 92; DB 20; Length 339;
Best Local Similarity 58.3%; Pred. No. 2.4e-13;
Matches 161; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 3464 CTTAGCAGATTTACAGACCTGCAGATGGGACACAGCAAGATGAACCATTTGACAGCCAGGA 3523
DB 12 CATGGCTACTTTGCGCACCGGACACACCCAGAAAGACGAGATTTGGAAATCCATGA 71
QY 3524 CAGTAAAGCCACTGCAGCTGTGAGGAGTCAAGGTTCAAGAGAGAGAGCGGCTACTGC 3583
DB 72 GGAGAATAGGTCGACTCTGGTACCAGTCAGGGGACAGAGAGCAGTCAGTCTCTGC 131
QY 3584 TCAGAAGAGGAGCCCTTCGACACTACTTAATGTTCCAGCCCAAGAGAAACATGGGA 3643
DB 132 ACAGAAAGAGAGGCGCTCCAGACCTTCCAGTTTGTGTTCCAGGAGAAACATAAGAAACA 191
QY 3644 AGAACCCAGGAAGAGATGTTCTTTGAACCTACACAGCAAGAGCTTACTGCTGCAGCCGTGCC 3703
```

Db 192 ATCAAAGATGGAAGACACTCTAGACATACACATTAAAGGAGTGTGAGTGAAGAACTGTATC 251
 QY 3704 CGTCTGGCAAGAAGACTGAGTGGGTCAAGAGGGTGA 3739
 Db 252 CATTCTGTCAAGAAGACTGAGGGGACTCAAGAGGCTGA 287

Search completed: December 12, 2002, 16:39:17
 Job time : 1217 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:13:30 ; Search time 166 Seconds
(without alignments)
11380.294 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cggcgccggggagtagaag.....aaataaacatatttgactac 6160

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4831.4	78.4	5134	2	US-08-635-121-1
2	2118.8	34.4	6605	1	US-08-769-309A-4
3	2118.8	34.4	6605	3	US-08-394-570-4
C 4	99.8	1.6	7218	1	US-08-232-463-14
C 5	71.2	1.2	1926	4	US-09-249-585A-4
C 6	71.2	1.2	1931	2	US-09-130-114-2
7	70.8	1.1	1926	4	US-09-249-585A-2
8	70.8	1.1	2580	3	US-09-050-863-2
9	70.8	1.1	2580	4	US-09-359-081-2
C 10	70.8	1.1	5452	2	US-09-130-114-1
11	70.8	1.1	9600	4	US-08-910-647-1
12	70.8	1.1	9600	4	US-09-620-925-1
13	70.8	1.1	10596	1	US-07-884-811-15
14	70.8	1.1	10596	1	US-07-885-971-15
15	70.8	1.1	10596	1	US-08-087-783A-15
16	70.8	1.1	10596	1	US-08-194-088B-15
17	70.8	1.1	10596	2	US-08-194-087-15
18	70.8	1.1	10596	5	PCT-US93-04648-15
19	66.4	1.1	3489	2	US-08-728-323A-1
20	66.4	1.1	3489	4	US-09-298-568-1
C 21	66.4	1.1	32207	2	US-08-770-379-20
C 22	66.4	1.1	32207	4	US-08-757-669A-20
C 23	66.4	1.1	32207	4	US-09-230-371A-20
C 24	64.6	1.0	16442	3	US-08-781-891-208
25	63.4	1.0	7218	1	US-08-232-463-14
26	60.4	1.0	2277	1	US-08-676-967-2
27	60.4	1.0	2277	1	US-08-676-974-2

28	60.4	1.0	2277	2	US-09-098-487-2	Sequence 2, Appli
29	60.2	1.0	3337	1	US-08-072-610-1	Sequence 1, Appli
30	60.2	1.0	3337	2	US-08-719-822B-1	Sequence 1, Appli
31	60.2	1.0	3337	4	US-09-092-458-1	Sequence 1, Appli
C 32	58.8	1.0	51259	3	US-08-781-891-209	Sequence 209, App
33	56.4	0.9	3211	2	US-08-574-959A-8	Sequence 8, Appli
34	56.4	0.9	3211	4	US-09-357-014-8	Sequence 8, Appli
35	56.4	0.9	3901	2	US-08-574-959A-6	Sequence 6, Appli
36	56.4	0.9	3901	4	US-09-357-014-6	Sequence 6, Appli
37	55	0.9	1052	1	US-08-466-603-1	Sequence 1, Appli
38	55	0.9	1052	1	US-08-314-503A-1	Sequence 1, Appli
39	55	0.9	1052	2	US-08-468-066-1	Sequence 1, Appli
40	55	0.9	1052	2	US-08-466-717-1	Sequence 1, Appli
41	55	0.9	1052	3	PCT-US95-12414-1	Sequence 1, Appli
42	55	0.9	1505	1	US-07-915-246-1	Sequence 1, Appli
43	52.6	0.9	1505	1	US-08-742-185-101	Sequence 101, App
44	51.8	0.8	43795	3	US-08-742-185-101	Sequence 101, App
45	47.8	0.8	2518	4	US-09-433-699-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-635-121-1
; Sequence 1, Application US/08635121
; Patent No. 5910442
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,121
; FILING DATE: 19-APRIL-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/33603
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5134 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-635-121-1

Query Match 78.4%; Score 4831.4; DB 2; Length 5134;
Best Local Similarity 97.9%; Pred. No. 0;

QY 3166 GGTCAAAGCACTGGAGAGGTTGAGGAGGTAGAGGAGGACTCCGAAGTGTGCTGCTCGG 3325
Db 2148 GGTCAAAGCACTGGAGAGGTTGAGGAGGTAGAGGAGGACTCCGAAGTGTGCTGCTCGG 2207
QY 3226 AGAAGAGAGGAGCTTTATGCGGAAAGGACCCGTGACAGGAGCTGGAGCTGAGCATCTTG 3285
Db 2208 AGAAGAGAGGAGCTTTATGCGGAAAGGACCCGTGACAGGAGCTGGAGCTGAGCATCTTG 2267
QY 3286 CACAGGCTCTGAGACTGGGACAGGCTACTCCAGAGAGCCCTTGAGTTCTTGAAGTCAAG 3345
Db 2268 CACAGGCTCTGAGACTGGGACAGGCTACTCCAGAGAGCCCTTGAGTTCTTGAAGTCAAG 2327
QY 3346 CAGATGTAGACCATGTCGCCACAGTCCAGGTTATCAAGCTCCAGCAGCTGATGAAACAG 3405
Db 2328 CAGATGTAGACCATGTCGCCACAGTCCAGGTTATCAAGCTCCAGCAGCTGATGAAACAG 2387
QY 3406 CCGTGGCCCTGAGTCAATCCGAAACCTTGACAGACAGTGAACAATGGAAGCACTCCCT 3465
Db 2388 CCGTGGCCCTGAGTCAATCCGAAACCTTGACAGACAGTGAACAATGGAAGCACTCCCT 2447
QY 3466 TAGCAGATTGACACACTGCGAGATGGGACACAGCAAGATGAAACCTTGACAGCCAGGACA 3525
Db 2448 TAGCAGATTGACACACTGCGAGATGGGACACAGCAAGATGAAACCTTGACAGCCAGGACA 2507
QY 3526 GTAAAGCCACTGACAGCTGTGAGGAGTCAAGGTGACAGAAAGAGCGGCTACTGCTC 3585
Db 2508 GTAAAGCCACTGACAGCTGTGAGGAGTCAAGGTGACAGAAAGAGCGGCTACTGCTC 2567
QY 3586 AGAAGAGGAGCTTCCACACTACTTAATGTTCCAGCCCGGAGAGACATGGGGAAG 3645
Db 2568 AGAAGAGGAGCTTCCACACTACTTAATGTTCCAGCCCGGAGAGACATGGGGAAG 2627
QY 3646 AACGAGGAAGAGATGTTCTTTGAACTACACAGCAAGAGCTTACTGCTGACAGCCGTGCCCG 3705
Db 2628 AACGAGGAAGAGATGTTCTTTGAACTACACAGCAAGAGCTTACTGCTGACAGCCGTGCCCG 2687
QY 3706 TTCTGGCAAGACTGAGGTGGGTCAAGAGGTGAGGTGACTGCTGGATGGAGAAAG 3765
Db 2688 TCTGGCAAGAGCTGAGGTGGGTCAAGAGGTGAGGTGACTGCTGGATGGAGAAAG 2747
QY 3766 TCAAAGAGAACAGGAGGTGTTGTACACTCTGACCCCAACAGTCAAAAGGCTGCTGATG 3825
Db 2748 TCAAAGAGAACAGGAGGTGTTGTACACTCTGACCCCAACAGTCAAAAGGCTGCTGATG 2807
QY 3826 TGACATATGACAGTGAAGTGAAGTGGAGTGGCCGGTGTGAGGAAAGGAGAGTACTGAAG 3885
Db 2808 TGACATATGACAGTGAAGTGAAGTGGAGTGGCCGGTGTGAGGAAAGGAGAGTACTGAAG 2867
QY 3886 TGCAGATCTTAGCTGGAGAGGAGAGATGGGAACTGACGTTGAAAGGAGGAAAGGG 3945
Db 2868 TGCAGATCTTAGCTGGAGAGGAGAGATGGGAACTGACGTTGAAAGGAGGAAAGGG 2927
QY 3946 AGACAAAGCCAGACAGTGAAGAGGTGAGCAGGAAACAGCCGCTCCTGAGCATG 4005
Db 2928 AGACAAAGCCAGACAGTGAAGAGGTGAGCAGGAAACAGCCGCTCCTGAGCATG 2987
QY 4006 AAGGAACTACGGAAGCAGTCTGACACTTGCATGTCATGTCAGCTCAGAGAGGGGAAGG 4065
Db 2988 AAGGAACTACGGAAGCAGTCTGACACTTGCATGTCATGTCAGCTCAGAGAGGGGAAGG 3047
QY 4066 CACTGGGAAGCCTTGGAGGAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGCAATAG 4125
Db 3048 CACTGGGAAGCCTTGGAGGAGCCCTTCTCTCCAGACCAAGACAAAGCAGGTTGCAATAG 3107
QY 4126 AGGTTCAAGTTCAAAGCTTGACACACAGTCACTCAAAACAGCAGAGCTGTGGAAGAG 4185
Db 3108 AGGTTCAAGTTCAAAGCTTGACACACAGTCACTCAAAACAGCAGAGCTGTGGAAGAG 3167
QY 4186 TCATAGAAACGGTTGTGATTTTACAGACAGGTGAAGTCCAGAGTGTGAGGTGCACACT 4245
Db 3168 TCATAGAAACGGTTGTGATTTTACAGACAGGTGAAGTCCAGAGTGTGAGGTGCACACT 3227
QY 4246 TATTACAGCTGAGAAAGTCTCTCTGCAACGGGTGGCCACTGGACTCTTTCAGCATGCAAGG 4305

Db 3228 TATTACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGGACTCTTCAGCATGCGAGAGG 3387
QY 4306 ACACGGTACCCCTGGGGCCCTGAGTCTCAGGCAGAAATCCCAATCATAGTAATCTCCTG 4365
Db 3288 ACACGGTACCCCTGGGGCCCTGAGTCTCAGGCAGAAATCCCAATCATAGTAATCTCCTG 3347
QY 4366 CTCTGAAAGACCCCTACATCTCTGACCTACAAGGAGAAATAAGCGCATCCAGAGAGAGC 4425
Db 3348 CTCTGAAAGACCCCTACATCTCTGACCTACAAGGAGAAATAAGCGCATCCAGAGAGAGC 3407
QY 4426 GATCAGAGGAGAGGACCAAGCCAGATGCTGCTCTGATGCTGACGCGNAGGAGATACAG 4485
Db 3408 GATCAGAGGAGAGGACCAAGCCAGATGCTGCTCTGATGCTGACGCGNAGGAGATACAG 3467
QY 4486 CAATCGAAAGAGTCTCAAGGCTGAAACCTGAGATCTCTGGAATCTTGAGAGTAAGAGCAACA 4545
Db 3468 CAATCGAAAGAGTCTCAAGGCTGAAACCTGAGATCTCTGGAATCTTGAGAGTAAGAGCAACA 3527
QY 4546 AGATTGTGTGAACGTCATTTCAGACAGCCGTTGACCAAGTTGCGACGTGACAGAAACAGCCC 4605
Db 3528 AGATTGTGTGAACGTCATTTCAGACAGCCGTTGACCAAGTTGCGACGTGACAGAAACAGCCC 3587
QY 4606 CCGAAACTCATGCTTATGATTTCAGACAGCCAGGTTCTCTGCAATGCGCTTGGACAGCAGGG 4665
Db 3588 CCGAAACTCATGCTTATGATTTCAGACAGCCAGGTTCTCTGCAATGCGCTTGGACAGCAGGG 3647
QY 4666 AGCCCAACAGATCTCGACAAATAATGAAAGATGCCAAGATGAAACACCCAGTCCGCGCAGC 4725
Db 3648 AGCCCAACAGATCTCGACAAATAATGAAAGTTCGCAAGATGAAACACCCAGTCCGCGCAGC 3707
QY 4726 CCAGAGAGACTTGCAGGCTCTGACCGTCTGAGGAGCATGGGCTCAGCCTCGGAAATGCT 4785
Db 3708 CCAGAGAGACTTGCAGGCTCTGACCGTCTGAGGAGCATGGCTCAG--CTCGGAAATGCT 3765
QY 4786 TGCAGGCTTGCAGTTGAAAGCCCGGTGTCAAAGTAAGCAATTTGAGAGAGTGTCTCTCA 4845
Db 3766 TGCAGGCTTGCAGTTGAAAGCCCGGTGTCAAAGTAAGCAATTTGAGAGAGTGTCTCTCA 3825
QY 4846 ACCCAAGAGTCCAAAAGGAGAGTGTGCTGATGCGCCCTCAGCTCCAAAGCTTTAGCCGAGG 4905
Db 3826 ACCCAAGAT--CAAAAGGAGAGTGTGCTGATGCGCCCTCAGCTCCAAAGCTTTAGCCGAGG 3884
QY 4906 CAGAGCCAGTGTCTCTGGAACCTTAACCAAGAAATCCCCAGACACCAACGAGCAAGC 4965
Db 3885 CAGAGCA--GTGTCTGGAACCTTAACCAAGAAATCCCCAGACACCAACGAGCAAGC 3941
QY 4966 TAAACGAGAGGGCGATCCCCCAAGAGTTGAGGTCCAGGAAGAAATGTTCTACCAAGT 5025
Db 3942 TAAACGAGAGCG--ATGCCCCCAAGAGTTGAGGTCCAGGAAGAAATGTTCTACCAAGT 4000
QY 5026 CAGTCAAGAGACCAAGCCCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5085
Db 4001 CAGTCAAGAGACCAAGCCCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4060
QY 5086 CAGAATCTTAAGATGTTAGTTGCTCAATGTATCATCTGTAAGACCAAGATGTGAAACAAG 5145
Db 4061 CAGAATCTTAAGATGTTAGTTGCTCAATGTATCATCTGTAAGACCAAGATGTGAAACAAG 4120
QY 5146 TCACAGAACAAAGATGCTGCTGTTGGACCTTGAGACCAAGATTTAGAGCCCATGAGATC 5205
Db 4121 TCACAGAACAAAGATGCTGCTGTTGGACCTTG--GACCAAGATTTAGAGCCCATGAGATC 4179
QY 5206 CAGAGAGCAGGGCCGCTCAATGATTTCCACCCAGTAGAGACCCCGACCAATTTCTGAGGCT 5265
Db 4180 CAGAGAGCAGGGCCGCTCAATGATTTCCACCCAGTAGAGACCCCGACCAATTTCTGAGGCT 4239
QY 5266 TCATCGGAGCTAGAGCAGCTAAACATTTCTCGTTTCAAGACTGCTTTGATTTGCCCC 5325
Db 4240 TCATCGGAGCTAGAGCAGCTAAACATTTCTCGTTTCAAGACTGCTTTGATTTGCCCC 4299
QY 5326 TTGATGCCGTCCGTGATTTTCTAAACATTTCTCGTTTCAAGACTGCTTTGATTTGCCCC 5385

Db 4300 TTGATGCCGTCGGTATTTTCAACATTTCTGTTTCAAGACGCTTTGATTTGCCCC 4359
 Qy 5386 TTGATGCCGTCGGTATTTTCAACATTTCTGTTTCAAGACGCTTTGATTTGCCCC 5445
 Db 4360 TTGATGCCGTCGGTATTTTCAAGACGCTTTGATTTGCCCC 4419
 Qy 5446 GGCATACCTAGTTCACCTCTCAAACTGGAGACCTCTTATGATTTATGATGCT 5505
 Db 4420 GGCATACCTAGTTCACCTCTCAAACTGGAGACCTCTTATGATTTATGATGCT 4479
 Qy 5506 TTTATGATGCTCTCTCTCTGTAACCTATTTATTTTCTTAAAGCTTTAAGCAAGCT 5565
 Db 4480 TTTATGATGCTCTCTCTCTGTAACCTATTTATTTTCTTAAAGCTTTAAGCAAGCT 4539
 Qy 5566 TTTTGTATTTATGCAATATTTAAGCGGTGTGCAAGCAAGCTTTTAAAGCTTTCCAA 5625
 Db 4540 TTTTGTATTTATGCAATATTTAAGCGGTGTGCAAGCAAGCTTTTAAAGCTTTCCAA 4599
 Qy 5626 GGCCTCAACGTAACTGCGACGAAAGATTAATCTCTGCGAAGAAAGACAAAGCTTTT 5685
 Db 4600 GGCCTCAACGTAACTGCGACGAAAGATTAATCTCTGCGAAGAAAGACAAAGCTTTT 4659
 Qy 5686 TTTAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGTGTGTTTCC 5745
 Db 4660 TTTAAAGTTTACTGATGCTTAGATCTGTGGGCTTCTAGTCTCTGAAAGTGTGTTTCC 4719
 Qy 5746 TATGACACGAGCTCAGAAATTAATAAATTTTGAACATCCAGATGTCCTCAATAT 5805
 Db 4720 TATGACACGAGCTCAGAAATTAATAAATTTTGAACATCCAGATGTCCTCAATAT 4779
 Qy 5806 TACCAATGATTTTCTCCCTTTTCTATGCAATGCAAGTGTGAAAGAGCTCTCTCT 5865
 Db 4780 TACCAATGATTTTCTCCCTTTTCTATGCAATGCAAGTGTGAAAGAGCTCTCTCTCT 4839
 Qy 5866 GTGTGATTAAGCCCTGCTCTTATGATGATGACAAATGAGTGTGCTTAAAGCCCATGA 5925
 Db 4840 GTGTGATTAAGCCCTGCTCTTATGATGATGACAAATGAGTGTGCTTAAAGCCCATGA 4899
 Qy 5926 GATGTTTCTTAATGCAAGAAATCTGTTGATGCTTTTGTATGATGCTCTTATGCT 5985
 Db 4900 GATGTTTCTTAATGCAAGAAATCTGTTGATGCTTTTGTATGATGCTCTTATGCT 4959
 Qy 5986 GGAACCGAATTCATGATGCAAGTGAAGTGAAGTCTGTTCTTTCAGATGATTTTGAATG 6045
 Db 4960 GGAACCGAATTCATGATGCAAGTGAAGTGAAGTCTGTTCTTTCAGATGATTTTGAATG 5019
 Qy 6046 ATACTGAGTTTGTCTGTATATCTGTGCTTCTTAAAGAAATGTTGATGATG 6105
 Db 5020 ATACTGAGTTTGTCTGTATATCTGTGCTTCTTAAAGAAATGTTGATGATG 5079
 Qy 6106 TTTCTTTGATTAATTTGATTTGACACTGATTTAAATTAACATATTTGACTTAC 6160
 Db 5080 TTTCTTTGATTAATTTGATTTGACACTGATTTAAATTAACATATTTGACTTAC 5134

RESULT 2
 US-08-769-309A-4
 ; Sequence 4, Application US/08769309A
 ; Patent No. 5741890

GENERAL INFORMATION:
 ; APPLICANT: Scott, John D.,
 ; APPLICANT: Naureit, Brian J.,
 ; APPLICANT: Kluack, Theresa M.
 ; TITLE OF INVENTION: Protein Binding Domains of Gravin
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower/233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/769,309A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: NO. 5741890and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 27866/33451
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-474-6300
 ; TELEFAX: 312-474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6605 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 192..5531
 ; US-08-769-309A-4

Query Match 34.4%; Score 2118.8; DB 1; Length 6605;
 Best Local Similarity 64.2%; Pred. No. 0;
 Matches 4206; Conservative 0; Mismatches 1802; Indels 546; Gaps 46;

Qy 3 GGGCCGCGGAGTGAAGAGCCACTGAGCCATGGGCGGAGGAGCTTCCAGGAGGCGG 62
 Db 162 GCGTATGCGCGGAGAAATGCGAGAGCCATGGGCGGAGGAGCTTCCAGGAGGCGG 221
 Qy 63 AGCCCGGAGAGCGG---GCGGGGAGCGACAGCGCGAGCTGTGCTCAGTGCCAT 119
 Db 222 AGCCCGGAGAGCGGCGCGGAGGAGCTTCCAGGCGGAGGCGGCGGCGG 281
 Qy 120 GGGCCGCGGAGTGAAGGCTCGGGA-----GCACTGAGAGACCCGCGGAGCGGAC 170
 Db 282 GGGCCCTGCGGCGGAGGCGGCGGCGGAGACACACCGGCGGAGCGGCGGCGGAC 341
 Qy 171 CCGCGCACCAAGCTCCACAGAAATGCGCAGCTGTCTTCTGTCACAGGCTAGCTAA 230
 Db 342 CCGCGCACCAAGCTCTTACAGAAATGCTCAGCTGTCCACCATCAATGAGGCTAG 401
 Qy 231 CAAGGAGTGTCCATGTCCAGAGGAAACCAGAGG----- 268
 Db 402 CAAGTGAAGTCAAGCTCAGAGGAGTGAATGAGCCAGAAAGAGCCCTGAAGGT 461
 Qy 269 -----GCAGAGAGAGAGTCTGTGATGAGATGTTGACAGCGA 308
 Db 462 CAAGAGCCCTTAACAGCGCAGAGAGAGAGAGTCAATGTTCAGAGAGTTGACAGAGA 521
 Qy 308 GAGTCAAGAGTGTGAG 368
 Db 522 GAGTCAAGAGTGTGAG 581
 Qy 369 GTTGAAGATCAAG 428
 Db 582 GTTCAAGATCAAG 638
 Qy 429 GCTTCAG 488
 Db 639 TCTTCAG 698
 Qy 489 GCTTCAG 548
 Db 699 GAGTTTAAG 758
 Qy 549 GAAAGTCAAGATGTCACACTCACTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 608

Db 759 GAGAAAGCTGACACTGCTCCAGCTACTCACTGTGAAGAAAGATGAAGGGAGGGAGCAGCA 818
Qy 609 GCCTCTGTGCGAGTGTGAGACACACAGAGCCCAAGTGTGAGACTGCGTGTGAGAGTCA 668
Db 819 -----GGGGCTGGGACACACAGGACCCAGCCTTGGGGCT-----GGAGAGCA 863
Qy 669 GCATCCAAAGAAAGTGTAGCTGAAGCAATCCACAGAGAAAGCAAGGACCCCTGAAGCAA 728
Db 864 GCATCCAAAGAAAGCAAGCAAGCAATCTACAGAGAAACCCGAAGAGACCCCTGAAGCGT 923
Qy 729 GAACAGAGCAGCAGAGAAATCCCTTCAAGCCGAATCTGATCAAGCGGTGAGAGAGAA 788
Db 924 GAGCAAGCCACGCGAGAAATTTCTCCCGCAGCCGAATCTGCGCAAGCAGT---GGAGGAA 980
Qy 789 GCCAAAGATGAAGGAGAGAAACCAAGAGAAAGAGCCCAAGTCCCAAGAAATCCCG 848
Db 981 TGCAAGAGGAGAGAGAGAGAAACCAAGAAAGAAACCTAGCAAGTCTGCAGAAATCTCCG 1040
Qy 849 AGCAGCCAGTCAACAGTGTAGACAAATCTTCTTCAAGAAAGTCTTCACTCAAGTGTGG 908
Db 1041 ACTAGTCCCGTGACAGTGTGAACAGGATCAACCTTCAAAAATTTCTTCACTCAAGTGTGG 1100
Qy 909 GCCGGCTGGGCAAGAGACAGCTTCAAGAAATCAAAAGAGGATGATCTGGAATCTGCC 968
Db 1101 GCCGGCTGGGCAAAAGACACAGTTTCAAGAACCGAGGAGGATGAAGTGAAGCTTCA 1160
Qy 969 GAGAAGAAAGGAGCAGAGCGAGAAAGTGTAGACAGAGGAAGAAAGAAAGACACAGAG 1028
Db 1161 GAGAAGAAAGGAAACAGAGCCAGAGCAAGAAAGTGTAGACACAGAGAAAGCGGAGAG 1220
Qy 1029 CCAGCCTCGGAGGA-----GCAGGAGCGCGCAGAGAC 1061
Db 1221 GTTGCCTCCGAGAAATGACCGCTCCGAGCAAGCCACACAGGAGCGCGCAGAAAGT 1280
Qy 1062 ACAGACGAGCCAGTGTGTAGCAGACTACAGAGAGTGTGAGCTGCCTTTGGAAGACAG 1121
Db 1281 GCCCAGAGCCCGGTATACAGTGTGAATATGAGAAAGTGTAGCTGCGCTTCAGAGAGCA 1340
Qy 1122 GTTGGTCACTGTGAGGAGT 1181
Db 1341 GTCAAGT 1400
Qy 1182 GATGAGAAAGTGAAGCCCAAGAA---GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1238
Db 1401 GATGAGAAATGAGAGTCCACCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1460
Qy 1239 GAGAAGACAGAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1298
Db 1461 GAGGAGAGAACCGAAGAGCAGAAA-----ACGGAGGTGGAA 1496
Qy 1299 GGAAACAGAGAAATCTTGTCCCTGTGAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1358
Db 1497 GAACACAGAGGTCTGT 1556
Qy 1359 GCTGAGCCTGT 1418
Db 1557 GCCCAACTGTGCAAGGAGT 1616
Qy 1419 ACTCAACTGACAGACCTTAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1478
Db 1617 ACACAGGAGGTGACCTCAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1676
Qy 1479 GTCAAGT 1538
Db 1677 GTGAGT 1736
Qy 1539 AAGAAACTTTTCACTAGT 1598
Db 1737 AAGAGCTTTTACAGCAGT 1796
Qy 1599 CGAGGAGGTGGGAGAGCAGAGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1658

Db 1797 AGAGGA---GGAGGAGACAGAGGAATCAGGGGAGCAGCACTCAGGTTCCAGCCGATTCTCCG 1853
Qy 1659 GAGAGTGTCTGATGAGCAGAGGAGGAGAGCTCTGCGTGTGCTCCCGAGGAGCCCTGAGGAG 1718
Db 1854 GACAGCCAGGAGGAGCAAAAGGCGGAGAGCTCTGCTCATCTCCCTGTAGGAGCCCGAGGAG 1913
Qy 1719 ACACAGTGTCTGAGAAAGGGCCGCTGTGAAGACACCCAGAGATGGGAGCTGTGAGGAAGGA 1778
Db 1914 ATCACGTGTCTGAAAAGGGCTTTAGCCGAGGTGACGAGATGGGAAGCTGTGAAGAAGGA 1973
Qy 1779 ACTACTTCCGATGAGAGAAAGAGAGAGAGAGGATCACTCCCTGGGCGATCTTCAAAAAG 1838
Db 1974 GCTACTTCCGATGAGAGAAAAAGAGAGAGTGTCTCTCCCTGGGCGATCATTCAAAAAG 2033
Qy 1839 ATGTTGACACCCCAAGAAACGGGTCCGAGACCTTTCTGAGAGTGACAAAGGAGGAGAGCTG 1898
Db 2034 ATGTTGACGCCCAAGAGCGTGTGTAGACGCCCTTCGAAAAGTGATAAAGAAAGATGAGCTG 2093
Qy 1899 GAGAAAGTCAAGAGCCGACCTTGTCTCTCACTGATAGCAGTGTCTGAGAAATGCAAGAT 1958
Db 2094 GACAAGGTCAAGAGCGCTACTTGTCTTCCACGAGAGCAGCCCTCTGAAATGCAAGAA 2153
Qy 1959 GAAAGTCAAAAAGCTGTTGTGTGAGGAACAAAAGCCAGAGGAACCAAGCGTAGGTGGATCT 2018
Db 2154 GAATGAAAGGGAGCGTGTGAAGAGCCAAAGCCGAAAGAACCAAGCCGCAAGGTGGATACC 2213
Qy 2019 TCAGTGTCTTTGGGAAGCAGCTGATTTGTGTGCGATCATCTCAAGAGAGAGAGCAAGGAGCA 2078
Db 2214 TCAGTATCTTTGGGAAGCTTTAATTTGTGTGGATCATCTCAAGAAAGAGCAAGGAGAGG 2273
Qy 2079 TCCTTTCAGATGATGAGAGGGCCAAAGGACACTGGGAGGGGACAGTCTCAGAGCAGAG 2138
Db 2274 TCCTTCTCTGATGAGGAAGGGGACCAAAAGCAATGTGGAGGAGACCAACCAAGAAAGCTGAT 2333
Qy 2139 GAGSCAGCAAGAACAAAGAGCGCGAAGACAGAGCTGTCTTCTGCCAGCACCCAGGAGCAG 2198
Db 2334 GAGSCCGGAAAAGACAAAGAGACGGGAGCAGAGCGGATCTTGTGTGTCTTCCAAAGAACAT 2393
Qy 2199 GACCAAGCGCAAGGAAGTCTTCTCACCCGAGCCAGCGGGAAGCCCTTCCGAGGGGGAAGGT 2258
Db 2394 GATCCAGGCGAGGAGTCTTCTCCCGAGCAGCTGTGAAAGCCCTACCGAAGGGGAGGCG 2453
Qy 2259 GTCTTCACTTTGGAGTCAATTTAAAAGATTAGTCACTCAAGAAAAAAATCAAGTCAAAA 2318
Db 2454 GTTTCCACTCTGGAGTCAATTTAAAAGTTAGTCAAGCAAGAAAAAAATCAAGTCCAAG 2513
Qy 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTTGTCCACTGAG 2366
Db 2514 CTGGAAGAGAAAGCGGAAGACTCCATAGCTGGGTCTGGTGTAGAACAATTCACATCCAGAC 2573
Qy 2367 ATCGAACCGAGTAGAGAAAGATCTTGGGTTTCCATTAAAGAAATTCATCCCGGAGCGCGG 2426
Db 2574 ACTGAACCCGTAAGAGAGATCTGGGTCTCAATCAAGAAATTTATCTTGGACGAGG 2633
Qy 2427 AAGAAAAAGGCGAGACGCGGAAGCAAGCAACTGTGTGAAGACTCAGGCGCAGTGGAG 2486
Db 2634 AAGAAAAAGGCGAGATGGGAACCAAGAACAGCCCTGTTGAAGACGCGAGGCGCAACAGGG 2693
Qy 2487 ATAAATGAGGACGACCTAATGTCCAGCCGTGTGCTCTGTGTGTGTGTGTGTGTGTGTGT 2546
Db 2694 GCCAACCAAGATGACTGTGATGTCCCGGCGTGTGCTCTGTGTGTGTGTGTGTGTGTGT 2753
Qy 2547 GAGAGGAGAGAGATGG-----AAGCCAGGGGAATACGGAGCTGCCCGCAGCTGCTG 2597
Db 2754 GAAAGGAGAGAAATGTGAGGACAGCAAGCCCAAAAGGCGCAGAGCAGCCGAGCAGAG 2813
Qy 2598 GGGGTGTGTGTCGTGTCCGAGGAGCTCAGTAAGACTCTGTGTCCACTGTGTGTGTGTGTGT 2657
Db 2814 GCAGCCACTCAGGTGTCCAAGGAGCTCAGCGAGAGTCAAGTGTGTGTGTGTGTGTGTGT 2873
Qy 2658 GTCAATGATGGGACGAGGCGAGTCAACAGTGTGTGAAGAGCGGTCTCTTGTGTGTGTGTGT 2717
Db 2874 GTGCTGTGACGGGACGAGGCGAGCTTACCATTATTGAAGAAAGGTCTCTTCTTGTGTGTGT 2933

Oy	2718	GCCTTCGTAACAAACCTCTTGAACACACAGCGGGAGAACCATGCACCTGTGAAAG	2777
Db	2934	GCTTCAGGACAGAAACCTCTTGAACAGTGAAGCTGAAGCCGCACTGTTACTGAAG	2993
Oy	2778	GTCACGTAAAAAGACATCATTCG---AGAGAAACTCTGTGCTCACCAAGCTTACCA	2834
Db	2994	GTAATTGGAAAGAAAGTAATTGCAAGAAAGAAACCCCCACGGTAACTGAACCTCTGCCA	3053
Oy	2835	GAGGTTAAAGTGGCCATTACACATGTGTCCACAGTGAAGTGTGATTTCACTCCGAAGCT	2894
Db	3054	GAGAACAGAGAGGCCCGGGCGCACAGCGTGTGTAAGAGCGAAATTGACCCCGAAGCT	3113
Oy	2895	GTGACAGCCACAGAGACCTCAGAGAGCTCTCCGTACTGAAGAAAGTTACCGAAGCATCGGG	2954
Db	3114	GTGACAGCTGCAGAAACTTGCAAGGGCCATTGGGTTCGGAAGAAAGAACGGAAGCATCTGCT	3173
Oy	2955	GCCGAAGAGACACACAGCATGTGTGTCCGACATTTCCAGCTGACTGATCTCCGACACCC	3014
Db	3174	GCTGAAGAGACCAAGAAATGTGTGACAGCTGTCCCAAGTTAACCGACTCCCGACACCC	3233
Oy	3015	ACAGAGGAAGCACCCCAAGTTCAAGAGGTGAAGAGGTGTGCTGATTCAGAAAGAAAG	3074
Db	3234	ACAGAGAGAGGCTACTCCGGTGCAGAGAGGTGAAGAGTGGCGTACCTGACATGAGAGGCA	3293
Oy	3075	GAGCGCCGAGACGACGGCCATCTCTCCAGCCGTTGCAGACAGAGGTGAAGAGAGTCCAG	3134
Db	3294	GAGAGGCGGACTCAAGAGGTCTTCAAGCAGTGGCAGAAAAAGTGAAGAGAAATCCAG	3353
Oy	3135	GTGCTTGCAACC---CAGACTGTGCAGAAACGGGGCTCAAAAGCACTGGAAGAAAGTTAG	3191
Db	3354	CTGCTGTGCACCGGTGGGCCAGAAAGATGTGCTTCACTGTGCAGAGACAGAGGCAAA	3413
Oy	3192	GAGGTAGAGAGAGACTCCGAAGTCTGGCTCCGAGAAAAAGAGAGCACTTATGCCGAA	3251
Db	3414	AGACCGAAGAGCAGGCTGAAGGCTCGGCTGTAAAGAAAGACGGAATGTAGTTGAA	3473
Oy	3252	GCACCCGTGCAGAAAGCTGAGCTGAGCATCTTGCAACAGGGCTGTGACATGCAAGGCT	3311
Db	3474	GTAGATGCTCAGAGAGGCAAAAACTGAGCCTTTTACCAAGAGGAAGGTGTGGGGCAGACC	3533
Oy	3312	ACTCCAGAGACCTCTG---AAGTTCCTGAAGTCAACGGCAGATGTAGAC-----CAT	3359
Db	3534	ACCCCAAGAAACTTTGAAAAAGCTCTTCAAGTCAACAGAGCATAGAGTCCAGTAGACTT	3593
Oy	3360	GTCCGACAGTCCCA-----GTTATCAAGCTCCAGCAGCTGTAGAA	3401
Db	3594	GTAACCACTTGTCAAGCCGAAACCTTAGTGGGGTAAAAATCACAGAGATGTGTAGAA	3653
Oy	3402	CAGGCGGTGGCCCTTGAGTCAATCCGAAACCTTGAACAGACAGTGAACAAATGGAAGCACT	3461
Db	3654	CAGGCTATCCCCCTGTGCTGGTGAAGAACCTTACAGACAGTGAAGCTGATGGAAGCACCC	3713
Oy	3462	CCCTTAGAGAAATTAGACACTGACAGATGGGACACACAAAGATGAACCAATTACAGCCAG	3521
Db	3714	CCCGTAGCCGACTTTGACGACACAGGACACCAAGAAAGAGAAATGTGGAATTCAT	3773
Oy	3522	GACAGTAAAGCCACTGCAGCTGTCAAGCACTCAAGGTCAAGAGAGAGGCGGCTACT	3581
Db	3774	GAGGAGAAATGCA---GGTGCACTTGTGTCCAGTCAAGGGGCAAGAAAGCAGGCACTTCT	3830
Oy	3582	GCTCAGAAAGAGAGCTTGCACACTACAAATAAATGTTCCAGGCCAGGAAGAAACATGG	3641
Db	3831	GCACGAAAGAGAGGCTCCACGACCTTCCAGTTTGTGTTCCAGGAAGAAACTTAAAGAA	3890
Oy	3642	GAAGAACACAGAAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTACTGCTCAGCCGTG	3701
Db	3891	CAATAAAGATGGAAGACACTCTAGACATATACGATTAAGAGAGGTGTCAATGGAACCTGA	3950
Oy	3702	CCCGTTCTGGCAAAAGCTGAGAGTGGGTCAAGAGGGTGA-----GGTTACGTGGTG	3752
Db	3951	TCCATTCTGTCAAAAGACTGAAGGGGAACTTCAAGAGGCTGACCAATATCTATTAAGAAAA	4010

Qy	3753	GAAGAGAAAAGTCAAAAGAAACAGAGAGGTGTTGACACTCTGG-----ACCCAAC	3806
Db	4011	AAAGACATACATTTTTCGAAGACTTGAGGGGCTTAATGACACAGGATPACAGTCACT	4070
Qy	3807	AGTCMAAAGGCTCTGATGTGACATATGACAGTGAAGATAGGGAGTGGCCGGGTCTCAG	3866
Db	4071	CGGAAAAGGTCACTGAAGTGTCCCTTAAGGTGAAGGACAGAAAGACGTGAATGTAAA	4130
Qy	3867	GAAAAGAGATGACTGAAGTGCAGAG-----TCTTAGCTTGGAGAGGGA	3911
Db	4131	AAGGATATCTCTTGAAGTCCAGAGTCAACGTAACTCTCCATCCCTCGTGGAGAGA	4190
Qy	3912	GAGATGGAACACTGACGTGAAAAAGGAGAAAGGGAGCAAGGCCAGAGCAAGTAGAGAA	3971
Db	4191	GAGATGTAGTTCAAGTCGAAGGGGAGAAAACAGAACAGAGCCACCCATGTAAAGAA	4250
Qy	3972	GAAGG---TGAGAGGAAACAGCCGCTCTCTGAGCATGGAAGAACCTACGGGAACCATC	4028
Db	4251	GAGAACTTGAGGACCGAAACAGCTGTTACCGTATCTGAAGAGGTCAATAGACACTCTCTC	4310
Qy	4029	CTGACACTTGAATGCECCAGCTCAGAGAGGGGAAAGCACTGGAAAGCTTGGAGAAC	4088
Db	4311	CAGACAGTGAATGTGCCATCATAGATGTGGGCAAGAAAGTACAGAGTTTGAAGAGAAC	4370
Qy	4089	CTTTCTC---TCCCAACCAAGCAAGAGGTGCTATAGAGTTCAAGTTCAAAGCTTG	4145
Db	4371	CTCTCTCTCCCTTGAAGTCAAGAGGAGGACGATATCACCAAAATTCAAGTTCAAGCTCT	4430
Qy	4146	GACACAAACGTCACTCAAACAGACAGAAAGCTGTGAAAA-----GGTCATGAAACGGTT	4199
Db	4431	GAGGATCATCTCACTCTTACACAGGCGCTGCAGAGAGAGAAAGGCTTTAGGAAAACTGCG	4490
Qy	4200	GTGATTTCAAGACAGGTGAATCTCAAGTGTGTAGTGCACATTATTAACAAGCTGAG	4259
Db	4491	AAACATTTAGAAACAGGTGAAACGTTGGAGCTGCAGAGGTGCACATTTAGTTCTGAAAGAG	4550
Qy	4260	AAATCTCTGCAACAGGTGGCCCACTGGAATCTTTCAGCATGACAGAGACACGGTACCCCTG	4319
Db	4551	AAATCTCTGTAAAAAATGAAAGCTTTGGCCGTCACTCAAGGGGAAAGATGCTGTGCCACA	4610
Qy	4320	GGGCGCTGAGTGCAGGAGAAATCATCCCATCATATGTAATCTCCGCTCGTGAAGACAC	4379
Db	4611	GGGCGGACCTGTCAAGGCAAAATTCACACAGTGAATGATCTGTCTACTTCCAAAGAAAGC	4670
Qy	4380	CTACATCTCTGACCTTCAAGAGAGAAATTAAGCGGATCCAGAGAGACGATCAGAGAAAG	4439
Db	4671	TTTAAGTCCGACCTGGAAGAGAGAAACACATCATCTGAAGTGAATCAGATGAAGTGC	4730
Qy	4440	GACAAAGCCAGATCTGTGCTCTGATGCTGACGGCAGAGAGAGTACAGCATGTAAAAAGTC	4499
Db	4731	GATGAGCAGGTGTGCTTGCCTC-----AGGAGGTCAAGTGAAGTGTAGCAATTTG---AGGAT	4781
Qy	4500	CTCAAGGCTGAACCTGAGATCTGGAACCTTGAAGTGAAGAGCAACAATGTGTCTGAC	4559
Db	4782	TTAGAGCTGAAAATGGGATTTTGGAACTTGAACCCAAAGGAGTAACTGTGCCAAAC	4841
Qy	4560	GTTCATTTCAAGCAGCCGTTGACAGTTTGCACGATAC---AGAAACAGCCCCGAAACTCAT	4616
Db	4842	ATCATCTCAAGACAGCCGTTGACAGATTTGTATCGTACAGAAAGAAACGCCAACCGAAATGTG	4901
Qy	4617	GCTTATGATTCACAGACCAGTTC-----CTGATSCAGAGCTTTCAG	4660
Db	4902	ACGTGTGATTTACAGACACAAAGCTCAACGTGATTAAGAGTGAACGACAGGACGCTGGACAG	4961
Qy	4661	CAGGAGCCCAACAGATGCTGGAACAAAAATGAAGATGCCAAGATGAACCCACATGCGC	4720
Db	4962	GAAGACGGAAGAAAGAGAGAAACCTCAAGGCTCTGCAACAGATGAAGAACCAATTAATCT	5021
Qy	4721	GGAGCCCA-----GAGAGCACTTGAAGCTCGACGGTCTTCTGGA	4759
Db	5022	TTAGCCAAAGAGAGATGACAGTACACCGCAGTGTGGCAAGCAACATTTCTGATATTTCCAAA	5081
Qy	4760	GGCATTGGGCTCAGCTCGGAATAGCTTTCGCGGCTTGCAGTTGAAGAGCGCGGTGTCAA-	4818

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 192..5531
US-08-994-570-4

Query Match 34.4%; Score 2118.8; DB 3; Length 6605;
Best Local Similarity 64.2%; Pred. No. 0;
Matches 4206; Conservative 0; Mismatches 1802; Indels 546; Gaps 46;

QY 3 GGGCCGCGGAGTGAAGAGCCATGACCATGGGCGAGGCAAGTTCCACCGAGAGAGGG 62
Db 162 GGCTAGGCGCGGAGAAATGCGAGAGACCATGGGCGCGGAGGCTCCACGAGAGCGC 221
QY 63 AGCCCGGAGAGCGCG---GCGGAGAGCAGACGCGGAGCGAGCTGGTCTCAGTGGCCAT 119
Db 222 AGCCCGGAGAGCGCGCGCGGAGGAGCTCCACGCGGCTGAGCCCGAGCCGAGCGCGG 281
QY 120 GGGCCGCGAGCTGAAGCCTCGGGA-----GCAGCTGAGAGCCCGCGAGCGGAGC 170
Db 282 GGGCCCTGCGCGAGGCGCGCGCAGACACACCGCGGACCCCGCATGCTGCTCGGAC 341
QY 171 CCGGCGACCAAGCTCCACAGAAAGATGGCCAGCTGTCTTCTGTCAAGGCGTAGCTGAA 230
Db 342 CCGGCGACCAAGCTCTTCAGAAAGATGCTCAGCTGCACCATCAATGGCGTAGCTGAG 401
QY 231 CAAGGAGATGTCATGTCCAGAGGAAACAGAGAGG----- 268
Db 402 CAAGATGAGCTCAGCTCCAGAGGAGGTGACCTAAATGGCCAGAAAGAGCCCTGAACGGT 461
QY 269 -----GCAGAGAGAAAGATGCTGTGATGAGAGATGTTGAGACGGA 308
Db 462 CAAGGAGCCCTAAACAGCAGAGAGAGAAAGATCATTTGTCAAGGAGTTGAGACAGAG 521
QY 309 GAGTGAAGATGTGAGAGAAAGAACCGAGTTGAAGAAATGGCGGCCAATCCACAGCT 368
Db 522 GACTGTGAAGATGTGAGAGAAAGAGACTCCGATAAAGATGGCTAATGACACGCGTT 581
QY 369 GTTGAAGATATCAACAAGATGGGAGAGAGAGCATCAAGAAATTAATTGAACAGATCCCT 428
Db 582 GTTCAGCATATCAGATGATGGGAGAGAGAG---ACCGAAATATGAGACAGATTCCT 638
QY 429 GCTTCAGAAACAATGTGAGAAAGATGATACAGCTGTGATGCCAGCTAATGATGTT 488
Db 639 TCTTCAGAAAGAAATTTGAAGAGCTAACAAACCCAGCTGATCCCAAGCTAATGATTT 698
QY 489 GCGTTCAAGAAAGTATTTAAATTTGTGTTTAAATTCACGCTGAAGAGATTAATAAT 548
Db 699 GGATTTAAGAAAGGTGTTTAAATTTGTGCTTTAAATTCACGTGTAATAAAGATTAAGCA 758
QY 549 GAAAGTGAATACTGCTCAACTACTACTGTCAAGAGAGATGAAGGCGAAGGCGCAGAA 608
Db 759 GAGAAAGCTGACATGCTTCACTACTGTGAGAAAGATGAAGGAGGAGGAGAGACAGCA 818
QY 609 GCGTCTGTGAGAGCTGAGACCAACAGAGAGCCAGTGTGAGAGACTGCCGTGAGAGATCA 668
Db 819 -----GGGGCTGGGAGCACACAGAGACCCAGCCTTTGGGGCT-----GGAAGAGA 863
QY 669 GCATCCAAAGAAAGTGAAGCTGAAGCAATCCACAGAAAGCAAGAGGACCTTGAAAGCA 728
Db 864 GCATCCAAAGAAAGGAAACCCAAACATCTACAGAGAAACCCGAAAGAGCCTTGAAAGCT 923
QY 729 GAACGAGCAGACAGCAAGAAATCCCTCTCAAGCGAATCTGATCAAGCGGCTGAGAGAA 788
Db 924 GAGCAAAAGCAGCAGAAATTTCTCCCGAGCGAATCTGGCCAAAGAGT---GAGAGAA 980
QY 789 GCCAAAGATGAAGAGAGAAAGAAACAGAGAAAGAGCCCAAGTCCCCAGATATCCCG 848
Db 981 TGCAAGAGAGAGAGAGAGAAACAGAGAAAGAAACCTAGCAATCTGCAAAATCTCCG 1040
QY 849 AGCAGCCAGTCAACAGTGAACAACATCTTCTTCAAGAAATCTTCACTCAGCGTTGG 908

Db 1041 ACTAGTCCCGTGACCAATGAAACAGATCAACCTTCAAAAATTTCTTCACTCAAGGTGG 1100
QY 909 GCCGCTGGGCGCAAGAAAGACCAAGCTTCAAGAAATCAAAAAGAGATGATCTGAAACTGCC 968
Db 1101 GCGGCTGGGCGCAAAAAGACCAAGTTCAGAGAGCCGAGAGATGAAGTGAAGCTTCA 1160
QY 969 GAGAAAGAAAGAGAGCAAGGCGAGAAAGATGACAGAGAAAGAAAGAAAGACAGAG 1028
Db 1161 GAGAAAGAAAGAGAAACAAGAGCCAGAAAGATGACACAGAAAGAGCGAAAGGCGAG 1220
QY 1029 CCAGCTCGAGAG-----GCAGAGAGCGGCGAGAGAC 1061
Db 1221 GTTCCCTCGAGAAATGACCCGCTCCGAGCAGACCCACCAAGAGCGGCGAGAAAGT 1280
QY 1062 ACAGACCAAGGCGAGTGTGACAGACTACAGAAAGGTGAGAGCTCTTGAAGACAG 1121
Db 1281 GCCACAGAGCCCGGTTATCAGCTGAATATGAGAAATGAGAGTGCCTCGAGAGAGCA 1340
QY 1122 GTTGTGACCTGAGAGCAATGCTCAGAGAGAGATGTCTCTTTGGCAAGAGTGT 1181
Db 1341 GTCAGTGCTCGAGAGAGCCTTGAAGAGAAACCTGCTCGTTGGCGACAGAGTGT 1400
QY 1182 GATGAGAAAGATGAAAGCCACCAAGAA---GTTGTGAGAGGTCAGAGAGACCGTG 1238
Db 1401 GATGAGAAATGAAAGTCCACAGAAAGAGTGTGGCCAAAGTCCAGTCAAGACCGTG 1460
QY 1239 GAGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGCTGAAGGGGCGTGTGTAGAA 1298
Db 1461 GAGAGAGAAACCGAAGAGCAGAA-----ACGAGAGTGA 1496
QY 1299 GGAACAGAGAAATCTTGGCCCTGAGAAACTGTGAGCCCAAGAGGTCCCCAGAA 1358
Db 1497 GAAACAGAGAGGTGTGCGCAGCTGAAGAAATGGTTGAAATGATGACAACTCAGAGAA 1556
QY 1359 GCTGAGCCTGCTGAGAGCTGATGAAGAGCAGAGATGTGTCTCTGAGAGAGACAC 1418
Db 1557 GCCGAACCTGCGAAGAGAGCTGTGATGACTCAAGAAACGTGTGTTCCGAGAGAGACCT 1616
QY 1419 ACTCACTGACAGACTTAAGTCTTGAAGAGAGACGCTGCCAAACCAAGAGAGATT 1478
Db 1617 ACACAGGAGAGTGAACCTCAAGCTGATGAGAGAGGTGTGTCCAAACCCCGAGAGGCTT 1676
QY 1479 GTCAGTGAGTGAAGATGCTGTCTTCAGAGAAAGATCAAGGTAAGAGAGTCCCTTG 1538
Db 1677 GTGAGTGAAGGTGAAGATGCTGTCACTACAGAGAGAAATGAAGGTGAGAGAGTCCACTA 1736
QY 1539 AAGAACTCTTCACTAGCTCAGAGCTTAAAGAGCTGTGGGAAAGAGCAAGAGGGGAAA 1598
Db 1737 AAGAACTTTTACAGACATGGCTTAAAGAGCTTTCTGAGAAAGAAACAGAAAGGGAAA 1796
QY 1599 CGAGAGGTGGGAGAGCAGAAAGCCTGAGAAATACCAACATTCACACCGAATCCCA 1658
Db 1797 AGAGAG---GGAGAGAGCAGAGAAATCAGGGGAGCACTCAGGTTCCAGCGCATTCG 1853
QY 1659 GAGAGTGTGATGAGCAGAGAGAGAGAGTCTGCGTGTCTCCCGAGAGAGCTGAGAG 1718
Db 1854 GACAGCAGAGAGAGCAAAAGGGGAGAGAGCTGTGCTCATCTCCCTGAGAGAGCCGAGAG 1913
QY 1719 ACCAGTGTCTGAGAAAGGGCGCTGGAAGCACCCAGAGATGGGGAAGCTGAGAAAGGA 1778
Db 1914 ATCAAGTGTCTGAGAAAGGGCTTAAAGGAGTGAAGAGAGAGAGAGTGAAGAGGA 1973
QY 1779 ACTACTTCAGTGAAGAGAGAAAGAGAGAGAGTCACTCCCTGGGAGATCTTCAAAAAG 1838
Db 1974 GCTACTTCAGTGAAGAGAGAGAGAGAGAGAGAGTCACTCCCTGGGAGATCTTCAAAAAG 2033
QY 1839 ATGTGACACCCAGAAAGAGGCTCCGAAAGACTTCTGAGAGTGAAGAGAGAGAGCTG 1898
Db 2034 ATGTGAGAGCCCAAGAGAGCGTTGAGAGAGCGCTTCGAGAAAGTGAATAAGAGAGAGCTG 2093
QY 1899 GAGAAAGTCAAGAGAGCGCACCTGTCTCCACTGATAGCACAGTGTCAAGAAATGCAGAT 1958

Dbb 2094 GACAAGGTCAGAGGGCTACCTTGTCTTCCACGAGAGCAGCCTCTGAAATCCAAAGAA 21153
Qy 1959 GAACTCAAAACTGTTGTGAGAAACAAAGCCAGAGAAACAAAGCGGTAGGTTGGATACT 2018
Dbb 2154 GAAATGAAAGGAGCGTGAAGAGCCAAAGCGGAAAGAACCAAGCGCAAGGTGGATACC 2213
Qy 2019 TCAGTGTCTTGGGAAGCACTGATTTGTGTCGGATCATCCAGAGAGAGAGCAAGGAAGCA 2078
Dbb 2214 TCAGTATCTTGGGAAGCTTTAAATTTGTGTGGGATCATCCAAAGAAAGAGCAAGGAAGG 2273
Qy 2079 TCCTCTTCAGATGATGAAGAGGGCCCAAGGACACTGGGAGGGGACAGTCAAGAGAGAGAG 2138
Dbb 2274 TCCTCTTCAGTGAAGAGGGGACCAAAAGCAATGGAGAGAGACCACCAAGAACTGAT 2333
Qy 2139 GAGGCCAGCAAGAACAAAGAACCCGGAACAGACGCTGTCTTCAGAGACCCAGAGAGAG 2198
Dbb 2334 GAGGCCGGAAGAACAAAGAGAGCGGGACAGACGGGATCCTTGTGTCTTCCAAAGAACAT 2393
Qy 2199 GACCAAGGCCAAGAAAGTTCTTCACTCCAGAGCCAGCGGGAAGCCCTTCGAAGGGGAAGGT 2258
Dbb 2394 GATCAGGGCAGGGAAGTTCTTCCCGAGCAAGCTGGAAGCCCTACCGAAGGGGAGGGC 2453
Qy 2259 GTCTCCACTTGGGAGTCATTTAAAGATTAAGTCACTCCAAAGAAAAAATCCAAGTCAAAA 2318
Dbb 2454 GTTCCACTGGGAGTCATTTAAAGGTTAGTCAAGCCAGAAAGAAATCAAGTCCAG 2513
Qy 2319 CTGGAAGAGAACCGGAGAC-----TCTAGTGTAGAGCAAGTTGTCCACTGAG 2366
Dbb 2514 CTGGAAGAGAAACGGAAGACTCCATAGCTGGGTCTGGTGTAGAACAATCCACTCCAGAC 2573
Qy 2367 ATCGAACCGAGTAGAAGAAATCTTGGTTTCCATTAAGAAATTCATCCCGGAGCGGG 2426
Dbb 2574 ACTGAACCGGTAAAGAAAGATCTCGGTCTCAATCAAGAAAGTTTATTTCTGGACGAAGG 2633
Qy 2427 AAGAAAAGGCGAGACGGAAGCAAGAACAGCCACTGTGGAAGACTCAGGSCCAGTGGAG 2486
Dbb 2634 AAGAAAAGCCAGATGGGAACAAGAACAGCCCTGTTGAAGACGAGGCGCCAAAGGG 2693
Qy 2487 ATAAATGAGGACACCTTAATGTCCAGCGGTGCTGCTGTCTGATATTAATGCAAGTG 2546
Dbb 2694 GCCAACGAAGATGACTGTGATGTCCCGCGGTGCTGCTGTCTGATGATGATGCTGTA 2753
Qy 2547 GAGGAGGAGAGATGG-----AAGCCAGGGGAATACGAGAGTGCCTCCAGCTGCTG 2597
Dbb 2754 GAAAGGAGAAATGGAAGGCAACAGAACAGAACAGCCCTGTTGAAGACGAGGCGCCAGCAAG 2813
Qy 2598 GGGGCTGTGACGTGTCGAGGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTGCGCA 2657
Dbb 2814 GCAGCCACTGAGTGTCAAGGAGCTCAGCGAGAGTCAGGTTTATATGATGGAGCAGCT 2873
Qy 2658 GTCAATTGATGGACAGGCGAGTCAACAGTGTGGAAGAGCGGTCTCCTTCTGATATACC 2717
Dbb 2874 GTCGCTGACGGAGGCGAGCTACCATTTATGAAGAAAGTCTCCTTCTTGGATATCT 2933
Qy 2718 GCTTCGTTACAGAACTCTTTGAACACACAGCGGGAGAGCCATGCCACCTGTTGAAGAG 2777
Dbb 2934 GCTTCAGTGACAGAACTCTTTGAACAAAGTAGAAGCTGAAAGCGGCACTGTTAACTGAGGAG 2993
Qy 2778 GTCACTAAAAGACATCATTCG---AGAGAACTCTGTGCTCACCAGAGCTTACCA 2834
Dbb 2994 GTATTGGAAGAGAAAGTAATTGAGAGAGAAACCCCGGTTACTGAACTCTGCCA 3053
Qy 2835 GAGGTTAAAGATGCCATGACGACATGCTCACCAGTGAAGTGAATTTCACTCAGAAAGCT 2894
Dbb 3054 GAGAAACAGAGAGGCCCGGGGCGACACGGTCTGTTAGTGAAGCGGAATTCACCCCGAAGCT 3113
Qy 2895 GTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAAAGTTACCGAAGCATCCGGG 2954
Dbb 3114 GTGACAGCTGCAGAAAACCTGCAAGGCGCATTTGGGTTCCGAAGAAAGGAAACCGAAGCATCTGCT 3173
Qy 2955 GCCGAAGAGACCAAGACATGTTGTCGCGAGTTTCCAGCTGACTGACTCCCGAGACACC 3014
Dbb 3174 GCTGAAGAGACCAAGAAAATGGTGTACAGCACTCTCCAGGTTAACCCGATTTCCCGAGACCC 3233

Qy 3015 ACAGAGAAAGCCACCCAGTTTCAGGAGGTAGAGAGTGGTGTCTGTAGATACAGAAAGAG 3074
Dbb 3234 ACAGAGAGGCCCACTCCGTTGCAGGAGGTGGAAGGTGGCGTACTCTGACATAGAAGACAA 3293
Qy 3075 GAGGCCAGACGACGAGCCATCTCCAGCCGTTTCGACACAGGTGGAAGAGGAGTCCCGAG 3134
Dbb 3294 GAGAGGCGGACTCAAGAGGTCTCCAGGCGAGTGGCAGAAAAAGTGAAGAGGAATCCCGAG 3353
Qy 3135 GTCCCTGCAACC---CAGACTGTGCAGAGAACGGGGTCAAAAGCACTGGAGAAAGTTGAG 3191
Dbb 3354 CTGCTGGCACCCGTTGGGCCAGAAAGATGTCTTCAGCCTGTGCAGAGAGCAGAGCCAGAA 3413
Qy 3192 GAGGTAGAGAGACTCCGAAGTGTGGCTTCGGAGAAAGAGAGAGAGCTTATGCCGAAA 3251
Dbb 3414 AGACCAGAAAGCAGGCTGAAGCGTGGGTCTGAAGAAAGAGACGGAATGTAGTGTGAAA 3473
Qy 3252 GGACCCGTGCAGGAAGCTGGAGCTGAGCATCTTCGACAGGGCTCTGAGACTGAGACAGGCT 3311
Dbb 3474 GTAGATGCTCAGGAGGCAAAACTGAGCCTTTTACAAAGGGAAGGTGGTGGGCGAGCC 3533
Qy 3312 ACTCCAGAGAGCCTTG---AAGTTCTGAACTCAGCGCAGATGTAGAC-----CAT 3359
Dbb 3534 ACCCCAGAAAGCTTTGAAAAGCTCTCAAGTCAAGAGAGCATAGAGTCCAGTGAGCTT 3593
Qy 3360 GTCCCAAGCTGCCA-----GGTTATCAAGCTCCAGCAGCTGATGGAA 3401
Dbb 3594 GTAAACCACTGTCAAGCCGAAACCTTAGCTGGGGTAAATCAACAGGAGATGGTATGGAA 3653
Qy 3402 CAGGCCGTGGCCCTGAGTCACTCCGAAACCTTCAGACAGCTGAGACAAATGGAAGCACT 3461
Dbb 3654 CAGGCTATCCCCCTGACTCGGTGGAAACCCCTACAGACAGTGAAGTGTGAGAGCACC 3713
Qy 3462 CCCTTAGCAGATTCAGACACTGCAGATGGGACACAGCAAGATGAACCATTTGACAGCCAG 3521
Dbb 3714 CCGTAGCCGACTTTGACGCAACAGGCAACAACCCAGAAAGACAGATTTGGAATCCAT 3773
Qy 3522 GACAGTAAAGCCACTGAGCTGTGAGCAGTCAAGGTCAAGAGAAAGAGCGGCTACT 3581
Dbb 3774 GAGGAGAAATGA---GGTGCATCTGGTACCACTCAGGCGGCACAGAGCAGAGGCTTCT 3830
Qy 3582 GCTCAGAAAGAGAGCCTTCGACACTACCTAATAATGTTCCAGCCGAGGAAGAATGGG 3641
Dbb 3831 GCACAGAAAGAGAGGCTCCAGACCTTCCAGTTTGTGTTCCAGGAAGAACTAAAGAA 3890
Qy 3642 GAAGAACCCAGGAAGAGATGTTTGAACCTACACAGCAAGACTTACTGTCGACGCGTG 3701
Dbb 3891 CAATCAAGATGGAAGACACTTAGAGCATACAGATAAGAGGTGTGAGTGAACCTGTA 3950
Qy 3702 CCCGTTCTGGCAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Dbb 3951 TCCATTCTGTCAAAGACTGAGGCGGACTCAAGAGGCTGACCAAGTATGCTGATGAGAAACC 4010
Qy 3753 GATGGAGAAAGTCAAGAAAGAACAGGAGGTGTTGTACACTCTGG-----ACCCAAC 3806
Dbb 4011 AAAGACGTACCAATTTTCGAAGGACTTGAAGGGTCTATAGACACAGGCATAACAGTCACT 4070
Qy 3807 AGTCAAAAGGCTGCTGATGTGACATATGACAGTGAAGTGTGGAGTGGCGGGTGTGAG 3866
Dbb 4071 CGGGAAGAGTCACTGAAGTTGCCCTTTAAAGGTGAAGGGGACAGAAAGCTGTAATGTAAA 4130
Qy 3867 GAAAAAGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGGAGGAGGGA 3911
Dbb 4131 AAGATGATGCTCTTGAACCTGCAGAGTCAAGCTTAAGTCTCTCTCCATCCCGTGGAGAGA 4190
Qy 3912 GAGATGAAACTGACGTTGAAAAGGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGTGA 3971
Dbb 4191 GAGATGTTAGTTCAAGTTCGAAAGGAGAGAAAACAGAAAGCAGAGCAACCCATGTGAATGAA 4250
Qy 3972 GAAGG---TGAGCAGGAACAGCGGCTCCTGAGCATCAAGGAACCTACGGGAAGCCAGTC 4028
Dbb 4251 GAGAAGCTTGAGCACGAAACAGCTGTTACCGTATCTGAAGAGGTGAGTAAGCAGCTCTC 4310


```
Db 6387 TGTAAG--TTTTGATCTACTCTTATATGCTGACTGCATTACACATGGCATGAATA 6444
QY 6014 AGTCCTGTTCTTACAGATGGTATTTGATAGATACTAGAGTTGTCTGTATATCTG 6073
Db 6445 AGTCAGGTTCTTTACAAATGGTATTTGATAGATACTAGGATTTGTGTGCCATATTG 6504
QY 6074 TGCCCTCTCTTAAAGACAATGTTGCAATATGTTCCCTTTCGATAAAATGTGATTGACAA 6133
Db 6505 TGCAATTCCTTTAAGACAATGTTGCAACATTCATTTCGATAAGTTGTGATTGACGA 6564
QY 6134 CTGATTTAAATAAA 6147
Db 6565 CTGATTTAAATAAA 6578

RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-F18
US-08-232-463-14

Query Match 1.6%; Score 99.8; DB 1; Length 7218;
Best Local Similarity 2.1%; Pred. No. 1.5e-15;
Matches 8; Conservative 262; Mismatches 109; Indels 0; Gaps 0;

QY 944 AAAAGAGGATGATCTGGAACCTGCCGAGAGAGAAAGGACGAGAGCAGAGAAAGTAGA 1003
Db 1431 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1372
QY 1004 CGAGGAAGAAAGAAAGACAGAGCCAGCTCGGAGGACGAGGACGGCGGACAGAC 1063
```

```
Db 1371 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1312
QY 1064 AGACCAAGCCAGGCTTGTACAGACACTACGAGAAGGTGGAGCTGCCTTTGGAACACCAAGT 1123
Db 1311 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1252
QY 1124 TGTGACCTGGAGGATCGTTCAGAGAGAGTGTGCTCTTTGGCAACGGAAGTGTGTA 1183
Db 1251 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1192
QY 1184 TGAGAAGATGGAGCCACCAAGAACTTGTGCAGAGGTCACGTCAGCACCGTCGGAGAA 1243
Db 1191 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1132
QY 1244 GACAGAGGAGGACGAGGAGGAGGAGGCTGAAGGGCGGTGTGTGTAGAGGAAC 1303
Db 1131 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1072
QY 1304 AGGAGAATCTTCCCTCCCT 1322
Db 1071 RRRRRATCGCAAGCTCCCT 1053

RESULT 5
US-09-249-585A-4/c
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match 1.2%; Score 71.2; DB 4; Length 1926;
Best Local Similarity 47.7%; Pred. No. 1.9e-08;
Matches 208; Conservative 0; Mismatches 228; Indels 0; Gaps 0;

QY 980 GGAGCAAGAGGCGCAGAAAAAGTAGACGAGGAAGAAAGAAAGACAGAGCAGCCTCGGA 1039
Db 977 GGAGGACGAGGACGGGGAGGACGAGGACGGGGAGGAGGACGAGGACGGGGACGGGA 918
QY 1040 GGAGCAGGAGCGCGCAGAGACACACAGCCAGGCTGTGTTCAGCAGACTACAGAGGTT 1099
Db 917 GGACGAGGACGGGGAGGACGGGGAGGACGAGGACGGGGAGGACGGGGAGGACGG 858
QY 1100 GGAGCTCCCTTTGGAAGACGAGTTGTGCTGACCTGGAGGACATCGTCAGAGGAGAGTGTGC 1159
Db 857 GGAGGACGGGGAGGAGGACGAGGACGGGGAGGAGGACGAGGACGGGGAGGAGGA 798
QY 1160 TCCTTTGGCAACCGAAAGTGTTCATGAGAAGATGGAAGCCACCAAGAAAGTTGTTCAGA 1219
Db 797 CGGGGAGGACGAGGACGGGGAGGACGGGGAGGACGAGGACGGGGAGGACGGGGAGGACGA 738
QY 1220 GTTCCACGTGAGCACCGTGGAGAGACAGAGGAGGACGAGGAGGAGGAGGAGGAGGCTGA 1279
Db 737 GGACGGGAGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGAGGACGGGGAGGACGG 678
QY 1280 AGGGGCGGTGTTGTAAGAGAAACAGAGAACTCTTTCCTCCCTGAGAAACTGGCTAGCC 1339
Db 677 GGAGGACGGGGAGGAGGACGAGGACGGGGAGGAGGACGAGGACGGGGAGGAGGACGG 618
```

[illegible]

RESULT 6
 US-09-130-114-2/c
 : Sequence 2, Application US/09130114
 : Patent No. 5976807
 : GENERAL INFORMATION:
 : APPLICANT: Horlick, Robert A.
 : APPLICANT: Damaj, Bassam B.
 : APPLICANT: Robbins, Alan K.
 : TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
 : TITLE OF INVENTION: From Multiple Transfected Episomes
 : FILE REFERENCE: 0867/19030US1
 : CURRENT APPLICATION NUMBER: US/09/130.114
 : CURRENT FILING DATE: 1998-08-06
 : NUMBER OF SEQ ID NOS: 36
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 2
 : LENGTH: 1931
 : TYPE: DNA
 : ORGANISM: EBNA
 : US-09-130-114-2

Query Match	1.2%;	Score 71.2;	DB 2;	Length 1931;
Best Local Similarity	47.7%;	Pred. No. 1.9e-08;		
Matches 208;	Conservative	0;	Mismatches 228;	Indels 0;
				Gaps 0;

QY	980	GGAGCAAGAGCCACAAAAAGTAAACGAGAAAGAAAAAGAAAAAGACAGAGCCACGAGCTCCGA	1039
Db	977	GGAGGACGAGGACCGGGAGGACGAGGACCGGGAGGAGGACGAGACCGGGAGGACCGGGAG	918
QY	1040	GGAGCAGAGACCGGCAGAGACACAGACCAAGCCAGAGTTGTCAAGACACTACGAGAGAGT	1099
Db	917	GGACGACGAGACGGGGAGGACCGGGGAGGACGAGGACCGGGGAGGACCGGGGAGGAGAGGACGG	858
QY	1100	GGAGCTGCTTTGGAAGCCAGGTTGGTGACTTGAGAGGACATCGTCAGAGAGAAATTGTGC	1159
Db	857	GGAGGACCGGGAGGAGAGGACGAGGACCGGGGAGGACCGGGGAGGAGACGAGGACCGGGAGGA	798
QY	1160	TCTCTTGGCAACGGAAGTGTTTGATGAAAGATGGAAGCCCAAGAAATTGTTGAGA	1219
Db	797	CGGGGAGGACCAAGGACCGGGGAGGACCGGGGAGGACGAGGACCGGGGAGGACCGGGGAGGACGA	738
QY	1220	GSTCCACCTGAGCACCTTGAGGAAGACAGAGGAGGACAGGAGGAGGAGGAGGAGGAGGCTGA	1279
Db	737	GGACCGGGAGGACGAGGACCGGGGAGGACGAGGACCGGGGAGGACGAGGACCGGGGAGGACGG	678
QY	1280	AGGGGCGTGGTGGTAGAAGAAAGAGAGATCCTTGCCCTCGAAGAAACTGGCTGAGCC	1339
Db	677	GGAGGACGGGAGGAGGAGACGAGGACCGGGGAGGACGAGGACCGGGGAGGACCGGGGAGGACGG	618
QY	1340	CCAGAGAGTCCCCACGAGAACTGAGCTTGAGGAGCTGATGAGAGCAAGAGATTGTG	1399
Db	617	GGAGGACGAGGACCGGGGAGGAGGACGAGGACCGGGGAGGACGGGGAGGACCGGGGAGGAGGA	558
QY	1400	TGTTCTTGAGGAGAC	1415
Db	557	CGAGGACCGGGGAGGAC	542

RESULT 7
US-09-249-585A-2
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:

```

? APPLICANT: Horlick, Robert
? TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
? FILE REFERENCE: 0867/0D905
? CURRENT APPLICATION NUMBER: US/09/249,585A
? CURRENT FILING DATE: 1999-02-11
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: patentin version 3.0
? SEQ ID NO 2
? LENGTH: 1926
? TYPE: DNA
? ORGANISM: Epstein Barr Virus
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1) ..(1926)
? OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

```

Query Match	1.1%;	Score 70.8;	DB 4;	Length 1926;
Best Local Similarity	45.3%;	Pred. No. 2.4e-08;		
Matches 335;	Conservative 0;	Mismatches 402;	Indels 3;	Gaps 2.

Qy	946	AAGAGGATGATCTCTGGAAATTCGCGAGGAAGAAAGAGCAAGGACAGAAAAATTAAGC	1005
Db	333	AGGAGGGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCGAGAGAGGGGGCAGG	332
Qy	1006	AGGAGAAAAAGAAAAAGACAGAGCCAGCCTCGAGAGACAGAGCCGCGACAGACACAG	1065
Db	393	AGCAGAGAGAGGGGGCAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGAGAGGGCGCAGAGCAGG	452
Qy	1066	ACCAAGGCCAGGTGTGTACAGAGACTTACGAGAGGTGGAGCTGCTTTTGGAGAACCAAGTTG	1125
Db	453	AGGAGGGGGCGAGAGGGGGCAGGAGCGAGAGAGAGGGGCGAGAGGGGCGAGAGAGC	512
Qy	1126	GTTGACTTGGAGGCATCGTCAAGAGAAATGTGTCTCTTTGGCAACCGAAATGTTTATG	1185
Db	513	AGGAGAGAGGGCGAGAGCAGAGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGAGGGGCGAGG	572
Qy	1186	AGGAAGATGGAGGCCCAACAGAAATTGTTCAGAGGTCCACGTGAGCACCGTGGAGAGA	1245
Db	573	AGGGGCGAGAGGGGCGAGAGCAGAGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGG	632
Qy	1246	CAGAGAGAGACAGAGGAGAGAGAGAGCTGTAAAGGGGGCTGTGTTAAGAAAGAACAG	1305
Db	633	AGGGGCGAGAGCA--GGAGGGGCGAGGACAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG	691
Qy	1306	GAGAAATCCTTGGCCCCCGAAGAACTGGCTGAGGCCCAAGAGGTCCCGAGAACTGAGC	1365
Db	692	GAGCA--GAGGGGCGAGAGCAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCG	749
Qy	1366	CTGCTGAGAGACTGATGATAGAGCAGAGAGATGTGTCTCTTGGAGAGACCACTCAAC	1425
Db	750	AGGAGGGGCGAGAGCAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGG	809
Qy	1426	TGACAGACTTAAGTCTTGAAGAGAGACGCTGCCCAACACCCAGAGAGCATTGTCAATG	1485
Db	810	AGCAGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGAGAGGGGCGAGAGGGGCGAGAGCAGG	869
Qy	1486	AGCTGGAATATCTCTCTCAGAGAAAGATCAAGGTACAGGAAAGTCCCTTGAAGAAC	1545
Db	870	AGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGG	929
Qy	1546	TCTTCAGTACGCTCAGGCTTTAAAGACTGTCTTGGAGAGAGCAGAGGGGAAACGAGAG	1605
Db	930	GGCAGAGACAGAGAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGTGAGAGGCCGGGG	989
Qy	1606	GTTGGGAGAGCAAGAGAGCTTGGAGAAATCCACACATTCAACCGAATCCCAAGAGATG	1665
Db	990	TGAGAGGAGCAGTGAAGGCCGGGGGTGAGAGGTTAGTGAAGGCCGGGGTTCGAGAGGTAG	1049
Qy	1666	CTGATGAGCAGAAAGGAGAG	1685
Db	1050	TGGAGGCCGCCGGGGTGAAG	1069

US-09-359-081-2

Query Match 1.1%, Score 70.8; DB 4; Length 2580;
Best Local Similarity 45.3%; Pred. No. 2.8e-08;
Matches 335; Conservative 0; Mismatches 402; Indels 3; Gaps 2;

QY 946 AAGAGATGATCTGGAAACTGCGGAGAGAAAGAGACCAAGAGCAAAAAGTGAACG 1005
DB 716 AGGAGAGGGGCGAGAGGGGCGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGG 775
QY 1006 AGGAG 1065
DB 776 AGGAG 835
QY 1066 ACCAGGCGAGGTTGTCAG 1125
DB 836 AGGAG 895
QY 1126 GTGACCTGAGAGCATCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
DB 896 AGGAG 955
QY 1186 AGAAGATGAG 1245
DB 956 AGGAG 1015
QY 1246 CAG 1305
DB 1016 AGGAG 1074
QY 1306 GAGAACTCTGCGCCCTGAG 1365
DB 1075 GAGCA--GGAGGGGAG 1132
QY 1366 CTGCTGAG 1425
DB 1133 AGGAG 1192
QY 1426 TGACAGACCTTAAGTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485
DB 1193 AGCAG 1252
QY 1486 AGCTGAGAGATGCTGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1545
DB 1253 AGGAG 1312
QY 1546 TCTTCAGTAGCTCAGGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1605
DB 1313 GGCAG 1372
QY 1606 GTGGGGAG 1665
DB 1373 TCGAG 1432
QY 1666 CTGATGAGCAG 1685
DB 1433 TCGAG 1452

RESULT 10
US-09-130-114-1/c
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassem B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/ID903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36

SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: YEENNA
US-09-130-114-1

Query Match 1.1%, Score 70.8; DB 2; Length 5452;
Best Local Similarity 45.3%; Pred. No. 4.3e-08;
Matches 335; Conservative 0; Mismatches 402; Indels 3; Gaps 2;

QY 946 AAGAGATGATCTGGAAACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1005
DB 2089 AGGAG 2030
QY 1006 AGGAG 1065
DB 2029 AGCAG 1970
QY 1066 ACCAGGCGAGGTTGTCAG 1125
DB 1969 AGGAG 1910
QY 1126 GTGACCTGAGAGCATCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
DB 1909 AGGAG 1850
QY 1186 AGAAGATGAG 1245
DB 1849 AGGAG 1790
QY 1246 CAG 1305
DB 1789 AGGAG 1731
QY 1306 GAGAACTCTGCGCCCTGAG 1365
DB 1730 GAGCA--GGAGGGGAG 1673
QY 1366 CTGCTGAG 1425
DB 1672 AGGAG 1613
QY 1426 TGACAGACCTTAAGTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485
DB 1612 AGCAG 1553
QY 1486 AGCTGAGAGATGCTGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1545
DB 1552 AGGAG 1493
QY 1546 TCTTCAGTAGCTCAGGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1605
DB 1492 GGCAG 1433
QY 1606 GTGGGGAG 1665
DB 1432 TCGAG 1373
QY 1666 CTGATGAGCAG 1685
DB 1372 TCGAG 1353

RESULT 11
US-08-910-647-1
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4

[illegible]

QY 1066 ACCAGGCCAGGTTGTCACAGACTACGAGAAAGGTGAGCTGCTTTGGAGACAGGTTG 1125
 DB 882 AGGAGGGGCGAGAGGGGCGAGAGCGAGAGAGGGGCGAGAGGGGCGAGAGGC 941
 QY 1126 GTGACCTGAGGCACTGTCAGAGAGAAAGTGTCTCTTTGGCAACGAAAGTGTGATG 1185
 DB 942 AGGAGAGGGGCGAGAGCGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGG 1001
 QY 1186 AGAAGTGAAGACCCACCAAGAAAGTTGTCAGAGGTCCACTGAGCACCTGTGAGAAAG 1245
 DB 1002 AGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGG 1061
 QY 1246 CAGAGGAGAGCGAGAGGAGAGAGAGGCTGAAAGGGGGCGTGGTGGTGAAGAGAACG 1305
 DB 1062 AGGGGCGAGAGCA-GGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 1120
 QY 1306 GAGAACTCTTGGCCCCCTGAGAACTGGCTGAGCCCCAGAGAGTCCCCAGAAAGCTGAGC 1365
 DB 1121 GAGCA--GAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCG 1178
 QY 1366 CTGCTGAGAGAGCTGATGAAGACAGAGAGAGTGTGTCTTGGAGAGACCACTCAAC 1425
 DB 1179 AGGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGG 1238
 QY 1426 TGACAGACTTAAGTCTTGAAGAGAACGCTGCCAACAACCCAGAGGCAATTTGCAATG 1485
 DB 1239 AGCAGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGG 1298
 QY 1486 AGGTGAGAGTGTCTCTCTCAAGAAAGATCAAGTACAGGAAAGTCCCTTGAAGAAAC 1545
 DB 1299 AGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGG 1358
 QY 1546 TCTTCAGTAGCTCAGGCTTAAAGAACTGTCTGGGAGAAAGCAAGAGGCAACGAGAG 1605
 DB 1359 GGCAGAGAGAGAGGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCG 1418
 QY 1606 GTGGGGGAGACGAAGGCTTGAAGAAATCAACACACTTACACCCGAATCCCAAGAGAGTG 1665
 DB 1419 TCAGAGAGCACTGAGAGGCGGGGCTCAGAGAGTAGTGAAGGCGGGGCTCAGAGAGGTAG 1478
 QY 1666 CTGATGACGAGAGAGGAGAG 1685
 DB 1479 TGGAGGCCGCGGGGTAGAG 1498

RESULT 13

US-07-884-811-15
 ; Sequence 15 Application US/07884811

; GENERAL INFORMATION:

; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
 ; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA

; ZIP: 94080
 ; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: patin (Genentech)
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/884,811

; FILING DATE: 19920518

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Ginger R.
 REGISTRATION NUMBER: 33,055
 REFERENCE/DOCKET NUMBER: 755.1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-3216
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10596 bases
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-07-884-811-15

Query Match 1.1%; Score 70.8; DB 1; Length 10596;
 Best Local Similarity 45.3%; Pred. No. 6.3e-08;
 Matches 335; Conservative 0; Mismatches 402; Indels 3; Gaps 2;

QY 946 AAGAGATGATCTGGAAGAACTCCGAGAAAGAAAGAGCAAGAGCCAGAAAGTAAAGC 1005
 DB 2296 AGGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGG 2355
 QY 1006 AGAAGAAAGAAAGAAAGCAAGAGCCAGCTTCGAGAGCGAGAGCCGCGAGAAAGACAG 1065
 DB 2356 AGCAGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGAGAGGGGCGAGAGGGGCGAGG 2415
 QY 1066 ACCAGCCAGTTGTTCAGACAGACTACGAGAAAGTGAAGTGTCTTTGAGAACCAAGTTG 1125
 DB 2416 AGGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGG 2475
 QY 1126 GTGACCTGAGAGCATCTGTCAGAGAGAAAGTGTCTCTTTGGCAACGAAAGTGTGATG 1185
 DB 2476 AGGAGAGGGGCGAGAGCGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGG 2535
 QY 1186 AGAAGATGAGAGCCCAACCAAGAGTTGTCAGAGATTCACGTAGACAACCTGAGAAAG 1245
 DB 2536 AGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGG 2595
 QY 1246 CAGAGAGAGAGAGAGAGAGAGAGAGCTTGAAGGGGCGTGGTGTGAAGAGAGAG 1305
 DB 2596 AGGGGCGAGAGCA-GAAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG 2654
 QY 1306 GAGAACTCTTCCCCCTGAGAACTGGCTGAGCCCCAGAGAGTCCCCAGAGAGCTGAGC 1365
 DB 2655 GAGCA--GAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGG 2712
 QY 1366 CTGCTGAGAGACTGATGAAGAGAGAGATGTGTCTCTGAGAGAGCACTCAAC 1425
 DB 2713 AGGAGGGGCGAGAGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGG 2772
 QY 1426 TGACAGACTTAAGTCTTGAAGAAAGACGCTGCCAACAACCCAGAGAGCAATTTGCAATG 1485
 DB 2773 AGCAGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGG 2832
 QY 1486 AGGTGAGATGTGTCTCTCAGAGAAAGATCAAGGTACAGGAAAGTCCCTTGAAGAAAC 1545
 DB 2833 AGGGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGG 2892
 QY 1546 TCTTCAGTAGCTCAGGCTTAAAGAACTGTCTGGGAGAAAGCGAGAGGAGAAACGAGAG 1605
 DB 2893 GGCAGAGAGAGAGAGAGGGGCGAGAGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGG 2952
 QY 1606 GTGGGGGAGACGAAGAGCTTGAAGAAATCAACACATTCAACCGAATCCCAAGAGAGTG 1665
 DB 2953 TCAGAGAGCACTGAGAGGCGGGGCTCGAGAGGTAGTGAAGGCGGGGCTCGAGAGGTAG 3012
 QY 1666 CTGATGACGAGAGAGAG 1685
 DB 3013 TGGAGGCCGCGGGGTAGAG 3032

RESULT 14

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 16:19:50 ; Search time 244 Seconds
(without alignments)
9983.770 Million cell updates/sec

Title: US-09-902-432-3
Perfect score: 6160
Sequence: 1 cgggcccggggagtagaag.....aaataaacattattgactac 6160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 355320 seqs, 19773052 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
10:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
12:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
14:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6160	100.0	6160	9	US-09-902-432-3
2	4831.4	78.4	5134	9	US-09-902-432-1
3	2144.4	34.8	6608	10	US-09-738-877-1
4	2144.4	34.8	6608	10	US-09-880-107-3439
5	1945.6	31.6	5346	10	US-09-738-877-2
6	550	8.9	563	10	US-09-917-800A-1223
7	162.2	2.6	178	10	US-09-917-800A-1475
8	106	1.7	720	9	US-10-101-487-74
9	106	1.7	720	9	US-10-101-487-76
10	89.6	1.5	522	9	US-10-101-487-71
11	89.6	1.5	530	9	US-10-101-487-73
12	89.6	1.5	554	9	US-10-101-487-69
13	89.6	1.5	554	9	US-10-101-487-106
14	71	1.2	575	10	US-09-864-761-20733
15	71	1.2	1969	10	US-09-864-761-3972
16	65.8	1.1	659158	9	US-09-771-208-20
17	64.4	1.0	100	10	US-09-728-445-336
18	61	1.0	700	10	US-09-864-761-17529
19	59.4	1.0	53226	10	US-09-818-264-3

ALIGNMENTS

RESULT 1

US-09-902-432-3
; Sequence 3, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6160
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-09-902-432-3

Query Match	100.0%	Score 6160;	DB 9;	Length 6160;
Best Local Similarity	100.0%	Pred No. 0;		
Matches 6160;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGGGCCGGGGAGTAGAAGAGCCACTGAGCCATGGCGGAGGAGTTCACCGAGCAGC	60	
Db	1	CGGGCCGGGGAGTAGAAGAGCCACTGAGCCATGGCGGAGGAGTTCACCGAGCAGC	60	
Qy	61	GGAGCCCGGAGCAGCCGGGGGAGCGACACCGCGAGGAGTGTCTAGTGGCCATG	120	
Db	61	GGAGCCCGGAGCAGCCGGGGGAGCGACACCGCGAGGAGTGTCTAGTGGCCATG	120	
Qy	121	GGCCCGGAGTGAAGCCCTCGGGAGCAGCTGGAGACCCCGCGAGCCGCCACCA	180	
Db	121	GGCCCGGAGTGAAGCCCTCGGGAGCAGCTGGAGACCCCGCGAGCCGCCACCA	180	
Qy	181	AGTCCACAGAAAGATGGCCAGCTGTTCTGTGTAACCGCGTAGCTAGTGAACAGGAGATG	240	

Sequence 22817, A
Sequence 6203, Ap
Sequence 3470, Ap
Sequence 1567, Ap
Sequence 20240, A
Sequence 27984, A
Sequence 11284, A
Sequence 68, Appl
Sequence 19707, A
Sequence 2926, Ap
Sequence 3, Appli
Sequence 7, Appli
Sequence 22, Appl
Sequence 4988, Ap
Sequence 11, Appl
Sequence 20, Appl
Sequence 3, Appli
Sequence 28274, A
Sequence 11696, A
Sequence 11856, Ap
Sequence 812, App
Sequence 3713, Ap
Sequence 20595, A
Sequence 11, Appl
Sequence 3, Appli
Sequence 52, Appl

|||||
Db 181 AGCTCCACAGAAAGATGGCCAGCTGTCTTGTCTACGGCTACTGAACAAGAGATG 240
Qy 241 TCCATGTCCAGAGAGAAACCAGAGGGGCAAGAGAAAGTCTGTATGAGATGTTG 300
Db 241 TCCATGTCCAGAGAGAAACCAGAGGGGCAAGAGAAAGTCTGTATGAGATGTTG 300
Qy 301 GACAGCCGAGACTCAGAGATGTGAGAGAAAAAGACCCAGTTGAGAAATGCGCGCAACT 360
Db 301 GACAGCCGAGACTCAGAGATGTGAGAGAAAAAGACCCAGTTGAGAAATGCGCGCAACT 360
Qy 361 CCAGAGCTGTGAGATATCAAAAGATGGGAGAGAGACATCAAAAATTAATTGAC 420
Db 361 CCAGAGCTGTGAGATATCAAAAGATGGGAGAGAGACATCAAAAATTAATTGAC 420
Qy 421 AGATCCCTGCTTCAGAAAAACAATGTGAGAAATGTGAACAGCTGCTGATGCCAGCTA 480
Db 421 AGATCCCTGCTTCAGAAAAACAATGTGAGAAATGTGAACAGCTGCTGATGCCAGCTA 480
Qy 481 ATGATGTTGCTTCAGAAAGTATTTAAATTTGTTTAAATTCAACGTTGAAGAG 540
Db 481 ATGATGTTGCTTCAGAAAGTATTTAAATTTGTTTAAATTCAACGTTGAAGAG 540
Qy 541 AATAAAATGAAGAAAGTCATCTGTCCAACTACTCACTCAAGAAAGATGAAGGCGAG 600
Db 541 AATAAAATGAAGAAAGTCATCTGTCCAACTACTCACTCAAGAAAGATGAAGGCGAG 600
Qy 601 GGGAGAAAGCTCTGTCCGAGCTGAGACCAACAGAGAGCCAGTGTGAGACTGCGCTG 660
Db 601 GGGAGAAAGCTCTGTCCGAGCTGAGACCAACAGAGAGCCAGTGTGAGACTGCGCTG 660
Qy 661 GAGAGTCAGCTCCAGAGAAAGTGAAGCAATCCACAGAGAAAGCAAGAGCACCC 720
Db 661 GAGAGTCAGCTCCAGAGAAAGTGAAGCAATCCACAGAGAAAGCAAGAGCACCC 720
Qy 721 TGAAGCAAGAACAGAGAGCAAGAAATCCCCCTTCAGCCGAATCTGATCAAGGGCTG 780
Db 721 TGAAGCAAGAACAGAGAGCAAGAAATCCCCCTTCAGCCGAATCTGATCAAGGGCTG 780
Qy 781 AGAAGAAAGCCAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 AGAAGAAAGCCAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 841 AATCCCGAGAGAGCCAGTCAACAGTGAACAATCTTCTTCAAGAGTTCTTCACTC 900
Db 841 AATCCCGAGAGAGCCAGTCAACAGTGAACAATCTTCTTCAAGAGTTCTTCACTC 900
Qy 901 ACGTTGGGCGGCTGGGGCAAGAGACCAAGCTTCAAGAAATCAAAAGAGATGATCTGG 960
Db 901 ACGTTGGGCGGCTGGGGCAAGAGACCAAGCTTCAAGAAATCAAAAGAGATGATCTGG 960
Qy 961 AAATGCGCGAG 1020
Db 961 AAATGCGCGAG 1020
Qy 1021 AGACAG 1080
Db 1021 AGACAG 1080
Qy 1081 CAG 1140
Db 1081 CAG 1140
Qy 1141 CGTCAG 1200
Db 1141 CGTCAG 1200
Qy 1201 ACCAAG 1260
Db 1201 ACCAAG 1260
Qy 1261 GAG 1320
Db 1261 GAG 1320

Db 1261 GAG 1320
Qy 1321 CTGAGAAACTGGCTGAG 1380
Db 1321 CTGAGAAACTGGCTGAG 1380
Qy 1381 TGAAG 1440
Db 1381 TGAAG 1440
Qy 1441 CTGAAG 1500
Db 1441 CTGAAG 1500
Qy 1501 CTTCTCAGAGAAAGATCAAGGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1501 CTTCTCAGAGAAAGATCAAGGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Qy 1561 GCTTAAAG 1620
Db 1561 GCTTAAAG 1620
Qy 1621 AGCTGAG 1680
Db 1621 AGCTGAG 1680
Qy 1681 GAGAGAGCTGCGCTGCTGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 GAGAGAGCTGCGCTGCTGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Qy 1741 GCGTGAAG 1800
Db 1741 GCGTGAAG 1800
Qy 1801 AG 1860
Db 1801 AG 1860
Qy 1861 TCCGAG 1920
Db 1861 TCCGAG 1920
Qy 1921 TGTCTCTCACTGATGAG 1980
Db 1921 TGTCTCTCACTGATGAG 1980
Qy 1981 AACAAAG 2040
Db 1981 AACAAAG 2040
Qy 2041 TTTGTGTGATCATCCAG 2100
Db 2041 TTTGTGTGATCATCCAG 2100
Qy 2101 GGGCAAG 2160
Db 2101 GGGCAAG 2160
Qy 2161 CCGGAG 2220
Db 2161 CCGGAG 2220
Qy 2221 CACCGAG 2280
Db 2221 CACCGAG 2280
Qy 2281 AAAGATTAGTACATCCAG 2340
Db 2281 AAAGATTAGTACATCCAG 2340
Qy 2341 CTAAGTGAAG 2400
Db 2341 CTAAGTGAAG 2400

QY 2401 TTAAGAAATTCATCCCGGACGGCGGAAGAAAGGGCAGACGGGAAGCAAGACCCA 2460
DB 2401 TTAAGAAATTCATCCCGGACGGCGGAAGAAAGGGCAGACGGGAAGCAAGACCCA 2460
QY 2461 CTGTGGAAGACTCAGGCGCAGTGGAGATAAATGAGGACGACCCCTTAATGTCCACGGCTCG 2520
DB 2461 CTGTGGAAGACTCAGGCGCAGTGGAGATAAATGAGGACGACCCCTTAATGTCCACGGCTCG 2520
QY 2521 TGCCTCTGTCTGATGATAATGCAAGTGGAGGGAGAGATGGAAGCCACGCGGGAATACGG 2580
DB 2521 TGCCTCTGTCTGATGATAATGCAAGTGGAGGGAGAGATGGAAGCCACGCGGGAATACGG 2580
QY 2581 AGCTGCCACAGCTGCTGGGGCTGTGTACGTGTCGAGGAGCTCAGTAAGACTCTGTCTC 2640
DB 2581 AGCTGCCACAGCTGCTGGGGCTGTGTACGTGTCGAGGAGCTCAGTAAGACTCTGTCTC 2640
QY 2641 ACACCTGTGAGTGTCCGAGTCAATTTGATGGGACACAGGGCAGTCCACAGTGTCCGAAGCGGT 2700
DB 2641 ACACCTGTGAGTGTCCGAGTCAATTTGATGGGACACAGGGCAGTCCACAGTGTCCGAAGCGGT 2700
QY 2701 CTCCTTCGTGGATATCCGCTTCGCTTAAACAGAACCTCTTGAACACACACGCGGGAGACCCA 2760
DB 2701 CTCCTTCGTGGATATCCGCTTCGCTTAAACAGAACCTCTTGAACACACACGCGGGAGACCCA 2760
QY 2761 TGCACCTGTGTAAGAGTCACTGAAAGACATCATTTGCAAGAACCTCTCTGTGCTCA 2820
DB 2761 TGCACCTGTGTAAGAGTCACTGAAAGACATCATTTGCAAGAACCTCTCTGTGCTCA 2820
QY 2821 CCACAGCTTTACAGAGGTAAAGATGCCATGACGACATGCTCACCAGTGTGAAAGTGGATT 2880
DB 2821 CCACAGCTTTACAGAGGTAAAGATGCCATGACGACATGCTCACCAGTGTGAAAGTGGATT 2880
QY 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCCGCTACTGAAAGAGTTA 2940
DB 2881 TCACCTCAGAAGCTGTGACAGCCACAGAGACCTCCGCTACTGAAAGAGTTA 2940
QY 2941 CCGAAGCATCGGGGGCCGAAGAGACACACATGTTGTCGAGTTTCCAGCTGACTG 3000
DB 2941 CCGAAGCATCGGGGGCCGAAGAGACACACATGTTGTCGAGTTTCCAGCTGACTG 3000
QY 3001 ACTCCACAGACACACAGAGGAAGCCACCCAGTTACAGGAGGTAGAGAGTGTGCTAG 3060
DB 3001 ACTCCACAGACACACAGAGGAAGCCACCCAGTTACAGGAGGTAGAGAGTGTGCTAG 3060
QY 3061 ATACAGAAGAGGAGCGCCAGACGCGAGCCATCTCTCCAGCCGTTGACAGCAAGGTGA 3120
DB 3061 ATACAGAAGAGGAGCGCCAGACGCGAGCCATCTCTCCAGCCGTTGACAGCAAGGTGA 3120
QY 3121 AAGAGAGTCCAGAGTGTGCAACCCAGACTGTGACAGAAACGGGGTCAAAAGCACTGG 3180
DB 3121 AAGAGAGTCCAGAGTGTGCAACCCAGACTGTGACAGAAACGGGGTCAAAAGCACTGG 3180
QY 3181 AGAAGGTTGAGGAGGTAGAGGAGCTCCAGTGTGCTGAGCATCTTGACAGGGCTCTGAGA 3240
DB 3181 AGAAGGTTGAGGAGGTAGAGGAGCTCCAGTGTGCTGAGCATCTTGACAGGGCTCTGAGA 3240
QY 3241 TTATGCGGAAAGACCCGTCGAGAGCTGGAGCTGAGCATCTTGACAGGGCTCTGAGA 3300
DB 3241 TTATGCGGAAAGACCCGTCGAGAGCTGGAGCTGAGCATCTTGACAGGGCTCTGAGA 3300
QY 3301 CTGGACAGGCTACTCCAGAGAGCTTTGAAGTTCCTGAAGTTCACGGCAGATGTAGACCATG 3360
DB 3301 CTGGACAGGCTACTCCAGAGAGCTTTGAAGTTCCTGAAGTTCACGGCAGATGTAGACCATG 3360
QY 3361 TCGCCAGCTCCAGGTTATCAAGTCCAGAGCTGATGGAACAGGCGGTGGCCCTGAGT 3420
DB 3361 TCGCCAGCTCCAGGTTATCAAGTCCAGAGCTGATGGAACAGGCGGTGGCCCTGAGT 3420
QY 3421 CATCCGAAACCTTTGACAGACAGTGAGACAAATGGAAGCACTCCCTTAGCAGATTTCAGACA 3480
DB 3421 CATCCGAAACCTTTGACAGACAGTGAGACAAATGGAAGCACTCCCTTAGCAGATTTCAGACA 3480

QY 3481 CTCAGATGGGACACAGCAAGATGAAACCATTTGACAGCCAGGACAGTAAGACCCACTGCGAG 3540
DB 3481 CTCAGATGGGACACAGCAAGATGAAACCATTTGACAGCCAGGACAGTAAGACCCACTGCGAG 3540
QY 3541 CTGTCAAGGAGTCAAGGCTCACAAGAAAGAGCGGCTACTGTCTCAGAAAGAGAGCCCTT 3600
DB 3541 CTGTCAAGGAGTCAAGGCTCACAAGAAAGAGCGGCTACTGTCTCAGAAAGAGAGCCCTT 3600
QY 3601 CGACACTACTAATAATGTTTCCAGCCCAGGAAGAACATGCGGGAAGAACCCAGGAAGAGATG 3660
DB 3601 CGACACTACTAATAATGTTTCCAGCCCAGGAAGAACATGCGGGAAGAACCCAGGAAGAGATG 3660
QY 3661 TTCTTTGAACTTACACAGCAAGAGCTTACTGCTCAGCGCTGCCCTTCTGCGAAAGACTG 3720
DB 3661 TTCTTTGAACTTACACAGCAAGAGCTTACTGCTCAGCGCTGCCCTTCTGCGAAAGACTG 3720
QY 3721 AGGTGGGTCAAGAGGTTGAGTTGACTGTTGGATGGAGAAAGTCAAAAGAAAGACAGG 3780
DB 3721 AGGTGGGTCAAGAGGTTGAGTTGACTGTTGGATGGAGAAAGTCAAAAGAAAGACAGG 3780
QY 3781 AGGTGTTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGCTGATGTGACATATGACAGTG 3840
DB 3781 AGGTGTTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGCTGATGTGACATATGACAGTG 3840
QY 3841 AAGTGTGGAGTGGCGGGTGTGAGAAAGAGAGATGCTGAAGTGCAGAGTCTTAGCC 3900
DB 3841 AAGTGTGGAGTGGCGGGTGTGAGAAAGAGAGATGCTGAAGTGCAGAGTCTTAGCC 3900
QY 3901 TGGAGAGGAGAGATGGAACCTGACGTTGAAAGAGAAAGGGAGACAAAGCCAGAGC 3960
DB 3901 TGGAGAGGAGAGATGGAACCTGACGTTGAAAGAGAAAGGGAGACAAAGCCAGAGC 3960
QY 3961 AAGTGTGAAAGAGTGGACGAGAAACAGCCGCTCTGAGCATGAAGAACCTTAGCGGA 4020
DB 3961 AAGTGTGAAAGAGTGGACGAGAAACAGCCGCTCTGAGCATGAAGAACCTTAGCGGA 4020
QY 4021 AGCCAGTCTGTGACACTGTCATGCCAGCTCAGAGAGGGGAAAGGCACTGGGAAGCCTTG 4080
DB 4021 AGCCAGTCTGTGACACTGTCATGCCAGCTCAGAGAGGGGAAAGGCACTGGGAAGCCTTG 4080
QY 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAGAGGTTGCAATAGAGTTCAAGTTCAAA 4140
DB 4081 GAGGAAGCCCTTCTCTCCAGACCAAGACAGAGGTTGCAATAGAGTTCAAGTTCAAA 4140
QY 4141 GCCTGGACACCAACAGTCACTCAAAACAGCAGAGAGCTGTGAAAGAGGTCATAGAAACGGTTG 4200
DB 4141 GCCTGGACACCAACAGTCACTCAAAACAGCAGAGAGCTGTGAAAGAGGTCATAGAAACGGTTG 4200
QY 4201 TGAATTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGACACTTTATTACAGCTGAGA 4260
DB 4201 TGAATTTCAGAGACAGGTGAAAGTCCAGAGTGTGTAGGTGACACTTTATTACAGCTGAGA 4260
QY 4261 AGTCCTCTCCAGCGGTGGCCACTGGACTCTTCAGCATGAGAGGACAGGTTACCCCTGG 4320
DB 4261 AGTCCTCTCCAGCGGTGGCCACTGGACTCTTCAGCATGAGAGGACAGGTTACCCCTGG 4320
QY 4321 GGCTCAGTCTCAGGAGAGTCCATCCCAATCATAGTAACTCTCTGCTCTGTAAGAGCACC 4380
DB 4321 GGCTCAGTCTCAGGAGAGTCCATCCCAATCATAGTAACTCTCTGCTCTGTAAGAGCACC 4380
QY 4381 TACATCTTGAACCTTACAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
DB 4381 TACATCTTGAACCTTACAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAGG 4440
QY 4441 ACAGCCAGATGCTGCTCTGATGCTGACGGCAAGAGAGATGACAGCAATGCAAAAAGTCC 4500
DB 4441 ACAGCCAGATGCTGCTCTGATGCTGACGGCAAGAGAGATGACAGCAATGCAAAAAGTCC 4500
QY 4501 TCAAGGCTCAACCTGAGATCTCTGGAACTTTGAGAGTGAAGAGCAACAGATTGCTCAAGC 4560
DB 4501 TCAAGGCTCAACCTGAGATCTCTGGAACTTTGAGAGTGAAGAGCAACAGATTGCTCAAGC 4560
QY 4561 TCATTGAGACAGCCGTTGACAGGTTGCGACGTACAGAAACAGCCCCGAAACTCATGCTT 4620

Db	4661	TCATTCACAGACCGCTTGAACCGATTCCGACGTCAGAAACAGCCCCGAAATCTATGCTT	4620
Oy	4621	ATGATTCACAGACCCAGGCTTCTGTCATGCAAGGCTTGAACAGAGGAGCCCAACAGATGCT	4680
Db	4621	ATGATTCACAGACCCAGGCTTCTGTCATGCAAGGCTTGAACAGAGGAGCCCAACAGATGCT	4680
Oy	4661	GCACAAAATGAAAGATGCCCAAGATGMAAACCCAGTGGCCGACGCCAAGAGGACTTGC	4740
Db	4661	GCACAAAATGAAAGATGCCCAAGATGMAAACCCAGTGGCCGACGCCAAGAGGACTTGC	4740
Oy	4741	AAGTCCTAACCGTTGCGAGGATGGGGCTCAGCTCCGAAATGCTTGCAGCGCTTGCACT	4800
Db	4741	AAGTCCTAACCGTTGCGAGGATGGGGCTCAGCTCCGAAATGCTTGCAGCGCTTGCACT	4800
Oy	4801	TGAAAGCCGCGGTCTCAAGTAAGCATTTAGAAGCTGCTCTTCAACCCAAAGATCAAA	4860
Db	4801	TGAAAGCCGCGGTCTCAAGTAAGCATTTAGAAGCTGCTCTTCAACCCAAAGATCAAA	4860
Oy	4861	AGGACATGCTGCTGATGAGCTCCAGCTCCAAAGCTTAAAGCTTAAAGCCAGAGGCGCCT	4920
Db	4861	AGGACATGCTGCTGATGAGCTCCAGCTCCAAAGCTTAAAGCTTAAAGCCAGAGGCGCCT	4920
Oy	4921	CTGGAACCTTAACCAAGATCCCAAGACCAACGGAACCAAGCTTAACGAGAGGCGC	4980
Db	4921	CTGGAACCTTAACCAAGATCCCAAGACCAACGGAACCAAGCTTAACGAGAGGCGC	4980
Oy	4981	ATCCCCCAAAATGTCAGTCTCAGGAAGAAATGTCTACAGTCAAGTCAAGAGAAC	5040
Db	4981	ATCCCCCAAAATGTCAGTCTCAGGAAGAAATGTCTACAGTCAAGTCAAGAGAAC	5040
Oy	5041	AGGCCAGGACAGAAAGAGACTGTCAGAGACCAAGGAGACTGTCAGAACTCTAAGATG	5100
Db	5041	AGGCCAGGACAGAAAGAGACTGTCAGAGACCAAGGAGACTGTCAGAACTCTAAGATG	5100
Oy	5101	TTAGTTGTCATTTGATCATCTGTGAACCAAGATGTGAAACCAAGTCACGACCAAGATG	5160
Db	5101	TTAGTTGTCATTTGATCATCTGTGAACCAAGATGTGAAACCAAGTCACGACCAAGATG	5160
Oy	5161	CTGCTGTGGAGCTTGGAGACCAAGATTTCAAGCCCATGAGTCAAGAGAGCGGCG	5220
Db	5161	CTGCTGTGGAGCTTGGAGACCAAGATTTCAAGCCCATGAGTCAAGAGAGCGGCG	5220
Oy	5221	TCCAAATGATTTCCACCCAGTAGACACCCCGACAAATTTCTAGGCTTCATGGGAGCTAGA	5280
Db	5221	TCCAAATGATTTCCACCCAGTAGACACCCCGACAAATTTCTAGGCTTCATGGGAGCTAGA	5280
Oy	5281	GCCAGCTTAACATTTCTCGTTTCAACATGCCCTTGAATTTGCCCTTATGCCGCTCGTG	5340
Db	5281	GCCAGCTTAACATTTCTCGTTTCAACATGCCCTTGAATTTGCCCTTATGCCGCTCGTG	5340
Oy	5341	TATTTCTAACATTTCTCGTTTCAAGACTGCTCTTGAATTTGCCCTTATGCCGCTCGTG	5400
Db	5341	TATTTCTAACATTTCTCGTTTCAAGACTGCTCTTGAATTTGCCCTTATGCCGCTCGTG	5400
Oy	5401	TATTTGGAATTAAGGTCCTGCTTCAACCTGGAACCAATTTCTGCATACCTAGTTCC	5460
Db	5401	TATTTGGAATTAAGGTCCTGCTTCAACCTGGAACCAATTTCTGCATACCTAGTTCC	5460
Oy	5461	ACTTCTCAAACTGAGACATCTCTTATATGATTTATATGATGTTTATAGTGTCTCC	5520
Db	5461	ACTTCTCAAACTGAGACATCTCTTATATGATTTATATGATGTTTATAGTGTCTCC	5520
Oy	5521	TCCGTAAACCAATGATTTATTTTCTTAAGCTTAAACATGCTTTTGTATATATGACAA	5580
Db	5521	TCCGTAAACCAATGATTTATTTTCTTAAGCTTAAACATGCTTTTGTATATATGACAA	5580
Oy	5581	TATATAACGGGTGACGCCATAGCGACTTGAAGAGCTCCAAAGCTCAACGTATAC	5640
Db	5581	TATATAACGGGTGACGCCATAGCGACTTGAAGAGCTCCAAAGCTCAACGTATAC	5640
Oy	5641	TGCAGCAACAGATACATTTCTGGCAAGAGACAAAGTCTTTTAAAGTTACGAT	5700

Db	5641	TGAGGAAAACAGTAAACATCTCCGCAAGAAAGACAAAGTCTTTTAAAGTTACGAT	5700
Qy	5701	GCTTAGATCTGTGGGCTTTACTCCTCTGAAAGGTGTGTTTCCCTAAGCACAGCGAGCT	5760
Db	5701	GCTTAGATCTGTGGGCTTTACTCCTCTGAAAGGTGTGTTTCCCTAAGCACAGCGAGCT	5760
Qy	5761	CAGAAATAAAAACCCCATTTTAAACATCCAGAGTGTCCCAATATTAACATGATTTTTC	5820
Db	5761	CAGAAATAAAAACCCCATTTTAAACATCCAGAGTGTCCCAATATTAACATGATTTTTC	5820
Qy	5821	CCCCCTTTTGCTAATCCAGTCCAGGTTGGAAAAGATCTCTCTGTGTCAGATTTAAGCC	5880
Db	5821	CCCCCTTTTGCTAATCCAGTCCAGGTTGGAAAAGATCTCTCTGTGTCAGATTTAAGCC	5880
Qy	5881	CTGTCTCTTAATGATATGACAAATAGAGTGTCCCTAAGCCCATGAGATGTTTCCTAATGC	5940
Db	5881	CTGTCTCTTAATGATATGACAAATAGAGTGTCCCTAAGCCCATGAGATGTTTCCTAATGC	5940
Qy	5941	AGAAAGATCTGTGTGACGTTTTTTTGATTTGATCTTTATGCTGTGACCGGAATTCATAT	6000
Db	5941	AGAAAGATCTGTGTGACGTTTTTTTGATTTGATCTTTATGCTGTGACCGGAATTCATAT	6000
Qy	6001	GCAATCGAAGTAGTACGCTGTTCTTTACAGATGATATTGTGTAGATACCTGAGATGTC	6060
Db	6001	GCAATCGAAGTAGTACGCTGTTCTTTACAGATGATATTGTGTAGATACCTGAGATGTC	6060
Qy	6061	TGTGTTATATCTGTGCCCTCTCTTAAAGAACATGTTGATTAATGTTCTCTTGGATTAAT	6120
Db	6061	TGTGTTATATCTGTGCCCTCTCTTAAAGAACATGTTGATTAATGTTCTCTTGGATTAAT	6120
Qy	6121	TGTGATTTGACAACTGATTTAAATTAACAATATTGACATAC	6160
Db	6121	TGTGATTTGACAACTGATTTAAATTAACAATATTGACATAC	6160

```

      RESULT 2
      US-09-902-432-1
        , Sequence 1, Application US/099024332
        , Patent No. US20020160002A1
        , GENERAL INFORMATION:
          , APPLICANT: Irwin H. Gelman
          , APPLICANT: Susan G. Jaken
            TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
            FILE REFERENCE: A30558-A-FWC-A 070156_0597
            CURRENT APPLICATION NUMBER: US/09/902,4332
            CURRENT FILING DATE: 2002-04-08
              PRIOR APPLICATION NUMBER: 08/978,277
                PRIOR FILING DATE: 1997-11-25
                  PRIOR APPLICATION NUMBER: 08/665,401
                    PRIOR FILING DATE: 1996-06-18
                      PRIOR APPLICATION NUMBER: 08/655,121
                        PRIOR FILING DATE: 1996-04-19
                          NUMBER OF SEQ ID NOS: 35
                            SOFTWARE: FastSeq for Windows Version 4.0
                              SEQ ID NO 1
                                LENGTH: 5134
                                  TYPE: DNA
                                    ORGANISM: Rattus norvegicus
                                      US-09-902-432-1

```

Query Match	78.4%	Score 4831.4	DB 9	Length 5134
Best Local Similarity	97.9%	Pred. NO. 0	Mismatches 76	Indels 31
Matches 5048	Conservative	0	Gaps	
QY	1016	GGAAAAAGCAGAGCCAGCCTCGAGAGAGCAGAGCCGCGCAGAAAGACACAGACCCAGGCCAG	1075	
Db	1	GGAAAAAGCAGAGCCAGCCTCGAGAGAGCAGAGCCGCGCAGAAAGACACAGACCCAGGCCAG	60	
QY	1076	GTGTGACAGACCTACGAGAAAGGTGAGCTGCTTTGGAAAGCCAGGTTGGTGAACCTCGGA	1135	
Db	61	GTGTGACAGACCTACGAGAAAGGTGAGCTGCTTTGGAAAGCCAGGTTGGTGAACCTCGGA	120	
QY	1136	GSCATCGTCAGAGGAAAGGTGCTCCTTTGGCAACGAAAGTTGTGATGAGAAGATGGA	1195	

Db	121	GGCATCGTCAGAGAGAAAGTGTGCTCCTTTGGCAACGGAAGGTGTTTGATGACGAAGATGGA	180
Qy	1196	AGCCCCACCAAGAAAGTTGTTGCAGAGGTCCACGTGAGCACCGTGGGAAAGACAGAGGAGGA	1255
Db	181	AGCCCCACCAAGAAAGTTGTTGCAGAGGTCCACGTGAGCACCGTGGGAAAGACAGAGGAGGA	240
Qy	1256	GCAGGGAGGAGGAGAGAGGCTGAAGGGGGCGTGTGGTGTAGAGAAACAGGAGAAATCCTT	1315
Db	241	GCAGGGAGGAGGAGAGAGGCTGAAGGGGGCGTGTGGTGTAGAGAAACAGGAGAAATCCTT	300
Qy	1316	GCCCCCTTGAGAAACTGSGCTCAGCCCCACAGAGGTCCCCAGGAAGCTGAGCCTGTGAGGA	1375
Db	301	GCCCCCTTGAGAAACTGSGCTCAGCCCCACAGAGGTCCCCAGGAAGCTGAGCCTGTGAGGA	360
Qy	1376	GCTGATCAAGAGCAGAGAGATGTGTCTCTGAGGAGACCACTCAACTCAGACACCT	1435
Db	361	GCTGATCAAGAGCAGAGAGATGTGTCTCTGAGGAGACCACTCAACTCAGACACCT	420
Qy	1436	AAGTCTTGAAGAGACGCTGCCAAAACACCCAGAGGCATTTGCTAGTGAGGTGGAGAT	1495
Db	421	AAGTCTTGAAGAGACGCTGCCAAAACACCCAGAGGCATTTGCTAGTGAGGTGGAGAT	480
Qy	1496	GCTGTCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTTGAAGAACTCTTCAGTAG	1555
Db	481	GCTGTCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTTGAAGAACTCTTCAGTAG	540
Qy	1556	CTCAGGCTTAAAGAGCTGTCTGGGAAGACAGAAAGGGGAAAACGAGAGGTGGGGGAGA	1615
Db	541	CTCAGGCTTAAAGAGCTGTCTGGGAAGACAGAAAGGGGAAAACGAGAGGTGGGGGAGA	600
Qy	1616	CGAAGAGCCTGGAGATACAAACACATTCACACCGAATCCCAGAGAGTGTGTATGAGCA	1675
Db	601	CGAAGAGCCTGGAGATACAAACACATTCACACCGAATCCCAGAGAGTGTGTATGAGCA	660
Qy	1676	GAAGGAGAGAGCTCTCGCTCGTCCCCCAGAGGCTTGAGGAGACCACTGTCTTGGAGAA	1735
Db	661	GAAGGAGAGAGCTCTCGCTCGTCCCCCAGAGGCTTGAGGAGACCACTGTCTTGGAGAA	720
Qy	1736	AGGCGCCTCGAAGACCCACAGGATGGGGAAGCTGAGGAAGGAATCTCTCCGATGGAGA	1795
Db	721	AGGCGCCTCGAAGAC-CCAGAGTGGGGAAGCTGAGNAGGNACTACTTC- - - - GTGGA	775
Qy	1796	GAAGAAGAGAGAAGGATCACTCCCTGGGCATCTTCTAAAAGATGGTGACACCCAAGAA	1855
Db	776	GAGAAGAAGAGGAAGGATCACTCCCTGGGCATCTTCTAAAAGATGGTGACACCCAAGAA	835
Qy	1856	ACGGTCCGAAGACCTTCTGAGAGTCAAAAGAGGAGAGCTGGAAGAGTCAAGAGCGC	1915
Db	836	AC- -GGTCCGAAGACCTTCTGAGAGTCAAAAGAGGAGAGCTGGAAGAGTCAAGAGCGC	894
Qy	1916	CACCTTCTCTCAGTAGACAGTGTCAAGATGCAAGATGAAGTCAAAATCTGTTGG	1975
Db	895	CACCTTCTCTCAGTAGACAGTGTCAAGATGCAAGATGAAGTCAAAATCTGTTGG	954
Qy	1976	TGAGGAACAAAGCCAGAGAACCAAGCGTAGGTGGATATCTCAGTGTCTTGGGAAGC	2035
Db	955	TGAGGAACAAAGCCAGAGAACCAAGCGTAGGTGGATATCTCAGTGTCTTGGGAAGC	1014
Qy	2036	ACTGATTTGTTCGGATCATCCAGAGAGACCAAGGAGGATCCCTCTCAGATGATGA	2095
Db	1015	ACTGATTTGTTCGGATCATCCAGAGAGAGCAAGGAGGATCCCTCTTCA- - -GATAT	1071
Qy	2096	AGGAGGCCAAGGACACTGGGA- -GGGGAACAGTCAAGAGCAGAGGAGGCCAGCAAGCA	2154
Db	1072	AGGAGGCCAAGGACACTGGGAGGGGAACAGTCAAGAGCAGAGGAGGCCAGCAAGCA	1131
Qy	2155	AAGAAGCCGGAACAGACGCTGTTCTGCCAGCACCCAGAGCAGGACCAAGCGCAAGAA	2214
Db	1132	AAGAAGCC- -GAAACAGACGCTGTTCTGCCAGCACCCAGGAGCAGGACCAAGCGCAAGAA	1190
Qy	2215	GTTCCTCACCCGACGAGCGGGAAGCCCTTCGAAGGGGAAGGTGCTTCCAATTGGGAGT	2274

QY 3346 CAGATGTAGACCATGTGTCCACAGTGTCCAGGTTATGAACTCCAGAGCTGATGGAACAG 3405
Db 2328 CAGATGTAGACCATGTGTCCACAGTGTCCAGGTTATCMACTCCAGAGCTGATGGAACAG 2387
QY 3406 CCGTGGCCCCGTGAGTCACTCCAAAACCTTGACAGACAGTGAACAAATGGAAGCACTCCCT 3465
Db 2388 CCGTGGCCCCGTGAGTCACTCCAAAACCTTGACAGACAGTGAACAAATGGAAGCACTCCCT 2447
QY 3466 TAGCAGATTGACAGACTGTCAGATGGGACACAGCAGATGAAACCAATTGACAGCCAGACA 3525
Db 2448 TAGCAGATTGACAGACTGTCAGATGGGACACAGCAGATGAAACCAATTGACAGCCAGACA 2507
QY 3526 GTAAAGCCACTGCACTGTCAAGGAGTCAAGGTCAAGAAAGAAAGGCGGTACTGCTC 3585
Db 2508 GTAAAGCCACTGCACTGTCAAGGAGTCAAGGTCAAGAAAGAAAGGCGGTACTGCTC 2567
QY 3566 AGAAAGAGAGCCTTCCGACCTAATTAATGTTCCAGCCAGGAAGAACATGGGGAAG 3645
Db 2568 AGAAAGAGAGCCTTCCGACCTAATTAATGTTCCAGCCAGGAAGAACATGGGGAAG 2627
QY 3646 AACCGAAGAGATGTTCTTGAACTTACACAGCAAGAGCTTACTGCTCAGCCGTGCCG 3705
Db 2628 AACCGAAGAGATGTTCTTGAACTTACACAGCAAGAGCTTACTGCTCAGCCGTGCCG 2687
QY 3706 TTTCTGGCAAGAAGCTGAGGTGGTCAAGAGGGTGAGGTTGACTGTTGGATGGAAGAAAAG 3765
Db 2688 TCTGGCAAGAAGCTGAGGTGGTCAAGAGGGTGAGGTTGACTGTTGGATGGAAGAAAAG 2747
QY 3766 TCAAGAAAGAACAGAGAGGTTTGTACACTCTGGAACCAAGCTCAAAAGGCTGCTGATG 3825
Db 2748 TCAAGAAAGAACAGAGAGGTTTGTACACTCTGGAACCAAGCTCAAAAGGCTGCTGATG 2807
QY 3826 TGACATATGACAGTGAAGTATGAGGAGTGGCCGGGTGTCAAGAAAAGAGATCTGAAG 3885
Db 2808 TGACATATGACAGTGAAGTATGAGGAGTGGCCGGGTGTCAAGAAAAGAGATCTGAAG 2867
QY 3886 TGCAGAGTCTTACCTCGAGAGGAGAGATGGAACCTGACCTTGAAGAAAGGAAAGGG 3945
Db 2868 TGCAGAGTCTTACCTCGAGAGGAGAGATGGAACCTGACCTTGAAGAAAGGAAAGGG 2927
QY 3946 AGACAAAGCCAGAGCAAGTGAAGAGTGAAGAGGAAACAGCCGCTCTGAGCATG 4005
Db 2928 AGACAAAGCCAGAGCAAGTGAAGAGTGAAGAGGAAACAGCCGCTCTGAGCATG 2987
QY 4006 AAGGAACCTACGGGAAGCCAGTCTGACACTTGACATGCCCCAGCTCAGAGAGGGGAAG 4065
Db 2988 AAGGAACCTACGGGAAGCCAGTCTGACACTTGACATGCCCCAGCTCAGAGAGGGGAAG 3047
QY 4066 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCATG 4125
Db 3048 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAAGACAAAGCATG 3107
QY 4126 AGGTTCAAGTTTAAAGCCCTGAGACACACAGTCACTTAAACACAGAAAGCTGAGAAAAG 4185
Db 3108 AGGTTCAAGTTTAAAGCCCTGAGACACACAGTCACTTAAACACAGAAAGCTGAGAAAAG 3167
QY 4166 TCATGAAAACGGTTGTGATTTAGAGACAGTGAAGAGTCCAGAGTGTAGAGTGCACACT 4245
Db 3168 TCATGAAAACGGTTGTGATTTAGAGACAGTGAAGAGTCCAGAGTGTGTAGGTGCACACT 3227
QY 4246 TATTTACCAAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGCACTCTTGACATGCAAG 4305
Db 3228 TATTTACCAAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGCACTCTTGACATGCAAG 3287
QY 4306 ACACGGTACCCCTGGGGCTGAGTCTCAGGCAAGATCCATCCCAACATAGTAATCTCTG 4365
Db 3288 ACACGGTACCCCTGGGGCTGAGTCTCAGGCAAGATCCATCCCAACATAGTAATCTCTG 3347
QY 4366 CTCCTGAAAAGCAACCTACATCCCTGACCTTCAAGGAAATTAAGCGCATCCCAAGAGAGC 4425
Db 3348 CTCCTGAAAAGCAACCTACATCCCTGACCTTCAAGGAAATTAAGCGCATCCCAAGAGAGC 3407

QY 4426 GATCAGAGGAAGAGACACAGCCAGATGCTGTCTGTAGTGCAGCGGACAGAGTACAG 4485
Db 3408 GATCAGAGGAAGAGAGACACAGCCAGATGCTGTCTGTAGTGCAGCGGAGAGTACAG 3467
QY 4486 CAATCGAAAAAGTCTTCAAGGCTTGAACCTGAGATCTGTGAACCTTGAGAGTAAAGACA 4545
Db 3468 CAATCGAAAAAGTCTTCAAGGCTTGAACCTGAGATCTGTGAACCTTGAGAGTAAAGACA 3527
QY 4546 AGATTGTCTGAACAGTCAATTCAGACAGCCGTTGACAGGTTGCACTGTACAGAAACAGCC 4605
Db 3528 AGATTGTCTGAACAGTCAATTCAGACAGCCGTTGACAGGTTGCACTGTACAGAAACAGCC 3587
QY 4606 CCGAACTCATGCTTATGATTCACAGACCAAGTCTGTGACATGCAAGGCTTGAACAGAGG 4665
Db 3588 CCGAACTCATGCTTATGATTCACAGACCAAGTCTGTGACAGGTTCTGCAATGCGCTTGAACAGAGG 3647
QY 4666 AGCCCAACAGATGCTGACAAAATGAAAGATGCAAGATGAAACACCCAGTGCAGC 4725
Db 3648 AGCCCAACAGATGCTGACAAAATGAAAGATGCAAGATGAAACACCCAGTGCAGC 3707
QY 4726 CCAAGAGAGATTTGCAAGTCTGACCGTTCTGAGAGCATGGGCTCAGCTTGGAAATGCT 4785
Db 3708 CCAAGAGAGATTTGCAAGTCTGACCGTTCTGAGAGCATGGGCTCAG--CTCGAAATGCT 3765
QY 4786 TGGCGGCTTGCAGTTGAAAGCGCGGTTGCAAGTGAAGCATTGAGAAAGTGCCTCTCA 4845
Db 3766 TGGCGGCTTGCAGTTGAAAGCGCGGTTGCAAGTGAAGCATTGAGAAAGTGCCTCTCA 3825
QY 4846 ACCCAAGATCCAAAAGAGAGCATGCTGATGGCCCTCAGCTCAAAAGCTTACCCAG 4905
Db 3826 ACCCAAGAT--CAAAAGAGATGCTGATGGCCCTCAGCTCAAAAGCTTACCCAG 3884
QY 4906 CAGAGCCAGTGCCTTGGAAAACCTTAACCAAAAGATCCCAAGACCAACGAGCAAAAGC 4965
Db 3885 CAGAGGCA---GTGTGGAACCTTAACCAAAAGATCCCAAGACCAACGAGCAAAAGC 3941
QY 4966 TAAACGAGAGAGGCGATCCCAAAAGTTGAGTCCAGAGAGAAAGAAATGCTACAGAT 5025
Db 3942 TAAACGAGAGAGG--ATGCCCAAAAGTTGAGTCCAGAGAGAAAGAAATGCTACAGAT 4000
QY 5026 CAGTCAAAAGAGACAAAGCCAGGCGAGAGAGACCTGCAAGAGCCAAAGGAGACCTGG 5085
Db 4001 CAGTCAAAAGAGACAAAGCCAGGCGAGAGAGACCTGCAAGAGCCAAAGGAGACCTGG 4060
QY 5086 CAGAACTCTTAAGTGTGTTGTTGCTCATTTGTATCTGTGTAACCAAGATGGAACAAAG 5145
Db 4061 CAGAACTCTTAAGTGTGTTGTTGCTCATTTGTATCTGTGTAACCAAGATGGAACAAAG 4120
QY 5146 TCACGAAACAAAGATGCTGCTGTGGGACCTTGAGACCAAGATTCAGAGCCCATGATC 5205
Db 4121 TCACGAAACAAAGATGCTGCTGTGGGACCTTGAGACCAAGATTCAGAGCCCATGATC 4179
QY 5206 CAGAGAGAGGCGCTCCCAATGATTTCCACCAGTAGAGCAACCCGACCAATTTGAGGCT 5265
Db 4180 CAGAGAGAGGCGCTCCCAATGATTTCCACCAGTAGAGCAACCCGACCAATTTGAGGCT 4239
QY 5266 TCATCGGAGCTAGAGCCAGCTAACATTTCTCGTTTCAAGCTGCTCTTGATTTGGCCC 5325
Db 4240 TCATCGGAGCTAGAGCCAGCTAACATTTCTCGTTTCAAGCTGCTCTTGATTTGGCCC 4299
QY 5326 TTGATGCGGCTCGGTATTTTCAACATTTCTCGTTTCAAGCTGCTCTTGATTTGGCCC 5385
Db 4300 TTGATGCGGCTCGGTATTTTCAACATTTCTCGTTTCAAGCTGCTCTTGATTTGGCCC 4359
QY 5386 TTGATGCGGCTCGGTATTTTCAAGCTGCTGCTGCTTCAACCTGGAACCAATTTCT 5445
Db 4360 TTGATGCGGCTCGGTATTTTCAAGCTGCTGCTGCTTCAACCTGGAACCAATTTCT 4419
QY 5446 GGCATACCTAGTTCATCTTCAAACTGAGAGATCTCTCTTATGATTTATGATGT 5505
Db 4420 GGCATACCTAGTTCATCTTCAAACTGAGAGATCTCTCTTATGATTTATGATGT 4479
QY 5506 TTTATGTAGTCTCTCTCTGTAACCTATTTGATTTTCTTCAAGTTTAAGCAATGCT 5565

QY 849 AGCAGCCGATCAACGATGAGCAACATCTTCTTCAAGAACTTCTCACTCAGCGTTGG 908
DB 1041 ACTAGTCCCGTGACCAAGTGAACAGATCAACCTTCAAAAAATTTCTCACTCAAGGTTGG 1100
QY 909 GCCGCTGCGCCCAAGAAACCAAGCTTCAAGAAATCAAAAAGAGATGATCTGGAACTGCC 968
DB 1101 GCCGCTGCGCCCAAGAAACCAAGCTTCAAGAAACCGAAGAGATGAAAGTGAAGCTTCA 1160
QY 969 GAGAAAGAAAGAGAGCAAGGAGCAAGAAAGTGAACGAGAAAGAAAGAGAAAGACAGAG 1028
DB 1161 GAGAAAGAAAGAGAGCAAGGAGCAAGAAAGTGAACAGAAAGAGAGAGAGAGAGAGAG 1220
QY 1029 CCAGCTTCGAGAG-----GACAGAGCCCGCAGAAAGAC 1061
DB 1221 GTTGGCTCCGAGAACTGACCGCTCCGAGCAAGCCACCCACAGAGCCGCGCAAGAGT 1280
QY 1062 ACAGACCAAGGCGAGTTGTCAGCAGACTACGAGAGTGAAGCTGCTTTGAGAACCGAG 1121
DB 1281 GCCACGAGCCCGGTTATCAGCTGAATGAGAAAGTTGAGCTGCCCTCAGAGAGAGCA 1340
QY 1122 GTTGTGACCTGAGAGCACTCGTCAAGAGAGAGTGTCTCTTTGCAACCGAGAGTGT 1181
DB 1341 GTCACTGGCTGCAAGGACCTTCTGAAAGAAACCTGCTCCGTTGGCGACAGAGAGT 1400
QY 1182 GATGAGAAAGTGAAGCCCAAGAA---GTTGTTGAGAGAGTCCAGTGAAGACCGTG 1238
DB 1401 GATGAGAAAGTGAAGTCCACCAAGAGAGTGTGTGCGAGAGTCCACCTCAGACCGTG 1460
QY 1239 GAGAGACAG 1298
DB 1461 GAG 1496
QY 1299 GGAACAGAGAGAACTCTTGGCCCTTGAGAACTGGCTGAGAGCCCGAGAGGTCCCGAGAA 1358
DB 1497 GAAACAGAGAGAGTGTGTGCTGACCTGAGAAATTTGTTGAAATGATGACAGAACTCAGAGAA 1556
QY 1359 GCTGAGCTCTGAGAGAGCTGATGAAGACAGAGAGATGTGTCTCTGAGAGAGACCAAC 1418
DB 1557 GCCGAACCTGACAGAGAGCTGTGTAAGCTCAAGAAAGAGAGAGAGAGAGAGAGAGAG 1616
QY 1419 ACTCAACTGACAGACCTTAAAGTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1478
DB 1617 AACAG 1676
QY 1479 GTCACTGAGTGAAGTGTCTCTCTCAGAGAAAGATCAAGGTACAGGAAATCCCTTG 1538
DB 1677 GTGAGTGAAGTGAAGTGTCTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1736
QY 1539 AAGAAACTCTTCAAGTACAGAGCTTTAAAGAGTGTCTGAGAGAGAGAGAGAGAGAGAG 1598
DB 1737 AAGAAAGCTTTTACAGAGCTGCTTTAAAGAGCTTTCTGAGAAAGAAAGAGAGAGAG 1796
QY 1599 CGAGAGAGTGGGAG 1658
DB 1797 AAG 1853
QY 1659 GAGAGTGTGATGAG 1718
DB 1854 GAG 1913
QY 1719 ACCAGCTGTCTGAG 1778
DB 1914 ATCAAGTGTCTGAG 1973
QY 1779 ACTACTTTCGATGAG 1838
DB 1974 GCTACTTTCGATGAG 2033
QY 1839 ATGAGTGAAG 1898
DB 2034 ATGAGTGAAG 2093

QY 1899 GAGAGAGTCAAG 1958
DB 2094 GAGAGAGTCAAG 2153
QY 1959 GAGTCAAAAGCTGTTGTGAG 2018
DB 2154 GAAATGAAG 2213
QY 2019 TCAGTGTCTTGGAG 2078
DB 2214 TCAGTGTCTTGGAG 2273
QY 2079 TCCTCTTCAGATGATGAAG 2138
DB 2274 TCCTCTTCAGATGATGAAG 2333
QY 2139 GAGGCGAGCAAG 2198
DB 2334 GAGGCGAGCAAG 2393
QY 2199 GACCAAGGCGAG 2258
DB 2394 GATCCAGGCGAG 2453
QY 2259 GTCTCACTTGGAG 2318
DB 2454 GTTTCACCTTGGAG 2513
QY 2319 CTGGAAG 2366
DB 2514 CTGGAAG 2573
QY 2367 ATCCAGCCGAGTGAAG 2426
DB 2574 ACTGAACCCGTTAAAG 2633
QY 2427 AAGAAAG 2486
DB 2634 AAGAAAG 2693
QY 2487 ATAAATGAG 2546
DB 2694 GCGAACGAG 2753
QY 2547 GAGAGGAG 2597
DB 2754 GAGAGGAG 2813
QY 2598 GGGGCTGTGTGAGTCCGAG 2657
DB 2814 GCGACCACTGAG 2873
QY 2658 GTCAATTGATGAG 2717
DB 2874 GTCCCTGACGGGAG 2933
QY 2718 GCTTCCGTTACAG 2777
DB 2934 GCTTACGATGAG 2993
QY 2778 GTCACTGAAAAAG 2834
DB 2994 GTATTGAG 3053
QY 2835 GAGGTTAAAGATGCCATGACAG 2894
DB 3054 GAG 3113
QY 2895 GTGACAGCCACAG 2954
DB 3114 GTGACAGCTGAG 3173
QY 2955 GCCGAAG 3014

Db 3174 GCTGAAGAGACCAAGAAATGGTGTAGCAGTCTCCAGTTAACCGACTCCCCAGACACC 3233
QY 3015 ACAGAGAAGCCACCCAGTTCAGGAGGTAGAGAGTGGTGTGTAGATACAGAGAAGAG 3074
Db 3234 ACAGAGAGGCCACTCCGGTGCAGGAGGTGAAGTGGCGTACTGCACATAGAAGACAA 3293
QY 3075 GAGGCCAGAGCAGGCCATCTCCAGCGGTTCGACAGAGTGAAGAGAGTCCAG 3134
Db 3294 GAGAGCGGACTCAAGAGGTCTCCAGGCCAGTGGCAAGAAAGTGAAGAGGAATCCAG 3353
QY 3135 GTCCCTGCAACC---CAGACTGTGCAGAGAAACGGGTCAAAAGCACTGGAGAAAGTTGAG 3191
Db 3354 CTGCTGGCACCGTGGCCAGAGATGTGCTTCAGCTGTGCAGAGAGCAGAGGCAGAA 3413
QY 3192 GAGGTAGAGGAGACTCCGAAGTGTGGCTTCGGAGAAAGAGAGCAAGTATATGCCGAAA 3251
Db 3414 AGACCAGAAGAGCAGGCTGAAGCTCGGGTCTGAAGAAAGAGACGGATGTAGTGTGAAA 3473
QY 3252 GGACCCGTGCAGGAAGCTGGAGCTGAGCATCTTGCACAGGGCTCTGAGACTGGACAGGCT 3311
Db 3474 GTAGATGCTCAGGAGGCAAAACTGAGCCCTTTACACAAGGGAAGTGGTGGGGCAGACC 3533
QY 3312 ACTCCAGAGAGCCTTG---AAGTTCTCTGAAGTCAAGCAGCATGTAGAC-----CAT 3359
Db 3534 ACCCAGAAAGCTTTGAAAAGCTCTCAAGTCAAGAGAGCATAGATCCAGTGAGCTT 3593
QY 3360 GTGCCACGTGCCA-----GGTTATCAAGCTCCAGCAGCTGATGGAA 3401
Db 3594 GTAACCACTTGTCAAGCGCAACCTTAGCTGGGGTAAATCAAGAGATGGTATGGAA 3653
QY 3402 CAGGCCGTGGCCCTGAGTCATCCGAAACCTTGACAGAGTGAACAATAAGAACACT 3461
Db 3654 CAGGCTATCCCCCTGACTCGGTGGAAACCCCTTACAGACAGTGAAGTGTATGGAAGCACC 3713
QY 3462 CCTTAGCAGATTACAGACTGCAGATGGGACACAGCAGAGTGAACCATTTGACAGCCAG 3521
Db 3714 CCGTAGCCGACTTGAGCCACAGGCAACACCAGAAAGCAGATTTGGAAATCCAT 3773
QY 3522 GACAGTAAAGCCACTGAGCTGTCCAGCAGTCAAGGTCAAGAAAGAGCGGCTACT 3581
Db 3774 GAGGAATGAGGTGCGCATCTGTGCCAGTCAAGGCGGCACAGAGAGCGAGTCTCT 3833
QY 3582 GCTCAGAAAGAGGAGCTTCGACACTACCTAATAATGTTCAGCCAGGAGAAACATGGG 3641
Db 3834 GCACAGAAAGAGAGGCTCCAGCACCTTCAGCTTTGTGTTCAGGAAGAACTAAAGAA 3893
QY 3642 GAAGAACAGGAAGAGATGTTCTTGAACCTACACAGCAGAGCTTACTGCTGAGCCGTG 3701
Db 3894 CAATCAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTGAGTGAACCTGTA 3953
QY 3702 CCCGTTCTGGCAAGACTGAGGTGGGTCAAGAGGGTGA-----GGTTGACTGGTTG 3752
Db 3954 TCCATTCTGTCAAGACTGAGGGACTCAAGAGGCTGACAGATGCTGATGAGAAACCC 4013
QY 3753 GATGGAGAAAAAGTCAAGAAAGAGGAGGTGTTGTACACTCTGG-----ACCCAAC 3806
Db 4014 AAAGAGTACCAATTTTTCGAAGACTTGAGGGTCTATAGACACAGGCATTAACAGTCACT 4073
QY 3807 AGTCAAAAGGCTGTGATGTGACATATGACGTGAAGTGAATGGAGTGGCCGGGTGTGAG 3866
Db 4074 CCGGAAAAGGTCACTGAAGTTCCTTAAAGGTGAAGGGACAGAAAGCTGAAATGTAAA 4133
QY 3867 GAAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCTGGAGGAGGA 3911
Db 4134 AAGGTATGCTCTTGAATGTGAGAGTCAAGTCAAGTCTCTCCATCCCTCCCGGTGGAGAGA 4193
QY 3912 GAGATGAAACTGACCTTGAAGGAGAGAAAGGAGACAAAGCAGAGCAAGTGAAGTGA 3971
Db 4194 GAGATGGTAGTTCAAGTCAAGAGGGAGAGAAAACAGAAAGCAGAGCCACCTATGTGAATGA 4253
QY 3972 GAAGG---TGAGCAGGAACAGCGCTCTCTGAGCATGAAGGAACCTTACGGGAGCCAGTGC 4028

Db 4254 GAGAGCTTGAGCACGAAACAGCTGTTACCGTATCTGAAGAGGTGAGTGAAGCAGTCTCTC 4313
QY 4029 CTGACACTTGCATGCCCCAGCTCAGAGAGGGGGAAGCACTGGGAAGCCTTGGAGGAAGC 4088
Db 4314 CAGACAGTGAATGTGCCCATCATAGATGGGCAAGAGTCAAGTGTGGNAGGAAGC 4373
QY 4089 CTTTCTC---TCCAGACCAAGACAAAGCAGGTGTGATAGAGTTCAGTGTCAAAGCCGTG 4145
Db 4374 CTTCTCTCCCTGCTAGGTCAAGAGGAGGAGTATGCACCAAAATTCAGTTCAGAGCTCT 4433
QY 4146 GACACAAAGTCACTCAAAACAGCAGAGCTGTGGAAA-----GGTCATAGNAACGGTT 4199
Db 4434 GAGGCATCATTTCACTCTAACAGCGCTGCAGAGGAGAAAAGTCTTAGAGAAAATGCC 4493
QY 4200 GTGATTTACAGACAGGTCAAAAGTCCAGAGTGTGTAGGTGCACACTTATATACCAGCTGAG 4259
Db 4494 AACATTTTAGAAACAGGTGAACGTTGGAGCTCGAGGTGCACATTTAGTTCTTGGAAAGAG 4553
QY 4260 AAGTCTCTGCAACCGGTGGCCACTGGGACTCTTTACAGCATGCAGAGGACACGGTACCCCTG 4319
Db 4554 AAATCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAAGATGCTGTGCCCA 4613
QY 4320 GGGCTGTAGTCTCAGSCAGAAATCCCAATCATAGTAACTCTCTCTCTGNAAGCACC 4379
Db 4614 GGGCCGACTGTCCAGCAAAATCGACACCAAGTGTAGTATCTGTACTACCAAGAAAGC 4673
QY 4380 CTACATCTCTGACTCAAGAGGAGAAATAAGCGCATCCAGAGAGAGCGATCAGAGGAAGAG 4439
Db 4674 TTAAGTTCCGACTCGAAGGAGAGAAACCATCATCTAGTGGNAGTCAAGTGAAGTC 4733
QY 4440 GACAAAGCAGATGCTGTGCTGATGTCAGGCAAGGAGAGTACAGCAATCGAAAAGTTC 4499
Db 4734 GATGAGCAGGTTGCTTGCC-----AGGAGGTCAAGTCAAGTGTGCAATTTG---AGGAT 4784
QY 4500 CTCAGGCTGAACCTGAGATCTCTGGAATCTTGAAGTGAAGCAACAGATTTGTCTGAAC 4559
Db 4785 TTAGAGCTGAAAAATGGGATTTTGAACCTTGAGACCAAAAGCAGTAACTTTGTCCAANAAC 4844
QY 4560 GTCAATTACAGACAGCGTTGACCCAGTTGCGACCTGAC---AGAAAACAGCCCCGAACTCAT 4616
Db 4845 ATCATCCAGACAGCGTTGACCCAGTTTGTACGTACAGAAAGAAACAGCCACGAAATGTTG 4904
QY 4617 GCTTATGATTCACAGACCCAGGTTCT-----CTGCATGCAGGCTTGACAG 4660
Db 4905 ACGTCTGAGTTACAGACACAAGCTCAAGTGTAAAGCTGACAGCCAGGACGCTGGACAG 4964
QY 4661 CAGGGAGCCCAACAGATGCTGCACAAAATGAAGATGCCAAGATGAACACCCAGTGC 4720
Db 4965 GAAACGGAGAAAGAGGAGAGGAACCTCAGGCTCTGCAACAGGATGAACACCAATTACT 5024
QY 4721 GCAGGCCA-----GAGAGACTTGCAAGTCTGACCGTTCTGGA 4759
Db 5025 TCAGCCAAAGAGGAGTCAAGTCAACCGCAGTGGGACAGACATTTCTGATATTCCAAA 5084
QY 4760 GGCATGGGCTCAGCCTCGGAAATGCTTGC CGGCTTGTCAGTTGAAAGCGCCGTTGTCAA- 4818
Db 5085 GACATGATGAAGCCTCAGAAAAGCCTGACTGTTGAGGTAGAGGTTCCACTGTAAAT 5144
QY 4819 ----- 4818
Db 5145 GATCAGCAGCTGGAAGAGTGTCTCTCCATCTGAGGAAGAGGAGGTGGAGCTGGAA 5204
QY 4819 -----AGTAAGCATTCAGAGCTGCCT 4840
Db 5205 AAGTCTGTGCCAAGAGATGATGCTCATGCTTTGTAGCAGAAAGAAATAGAGAAAGTCACTA 5264
QY 4841 CTTCAACCCAAAGA---TCCAAAAGGAGCATGCTGTGATGGCCCTCAGCTCCAAAGCTTA 4898
Db 5265 GTTGAACCGAAGAGATGAAGAAAGGTGATGATGATGACCTTGAAAACACAGAACTCA 5324
QY 4899 GCCCAGGAGGAGGCGCAGTGTCTCTGGAACCTTAACCAAGAAATTCCTCCAGACACCAACCGA 4958
Db 5325 GCGCTGGCTGACTGATGCTGCTCAGGAGGCTTAAACCAAGAGTCCCCAGATACAAATGGA 5384

```

QY 4959 CCAAGCTAACCGAGGAGGCGGATCCCCCAAAAGTTGAGTCCAGAGAGAAATGCT 5018
Db 5385 CCAAAAGAGAGAGAGAGAGAGATGCCAGAGAGATGAGATTCAGAGAGAGAAATGAC 5444
QY 5019 ACCAGTACGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5078
Db 5445 AGTGAATAGATTAAGAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5504
QY 5079 GACCTGGAGAGATCTTAAGATGTTAGTTG-----CTCATTGAC 5117
Db 5505 GAATGCGCAAGTGCAGACTTAACAGATCTTAAACATGACAGTTAACTCATTTGCT 5564
QY 5118 ATCTGTAAGACCAAGATGTAAGAAACAGATCAGAGAGAGAGAGAGAGAGAGAGAGAG 5173
Db 5565 GTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5623
QY 5174 CTGAGAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5233
Db 5624 CTGAAGAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5682
QY 5234 ACCAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5293
Db 5683 TTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5741
QY 5294 TCCCTGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5353
Db 5742 TCCCTGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5781
QY 5354 TCCCTGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5413
Db 5782 -----CATTAATTCGATTTA 5799
QY 5414 AGGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5472
Db 5800 AGGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5859
QY 5473 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5532
Db 5860 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5916
QY 5533 TGTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5575
Db 5917 TGTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5976
QY 5576 TGCATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5635
Db 5977 CACAGATATATATATATATATATATATATATATATATATATATATATATATATATAT 6036
QY 5636 TAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5686
Db 6037 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6096
QY 5687 TAAAGTTATATATATATATATATATATATATATATATATATATATATATATATATAT 5746
Db 6097 TAAATTTACTATATATATATATATATATATATATATATATATATATATATATATATATAT 6156
QY 5747 ATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5806
Db 6157 ATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6216
QY 5807 ACCATGATTTTTTCCCTTTTCTTATATATATATATATATATATATATATATATATATAT 5866
Db 6217 ATTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6269
QY 5867 TGTACAG-----TTAAGCCGTGCTCTTAATGATATGAGCAATAGT 5909
Db 6270 GGTAGCAAAATTAATCTTACATCTTTTCTTACTTTTATGTTGTTTGAACCAATAGT 6329
QY 5910 GTGCTTAAGGCGCATGAG-----TGTTTCTTAATGAGAGAGAGAGAGAGAGAGAGAG 5973
Db 6330 GTGCTTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6389

```

```

QY 5954 TGTAGTTTTTTTGTATGTAATCTCTTCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAG 6013
Db 6390 TGTAG-----TTTTGATTTCTACTCTTATATATGCTGAGAGAGAGAGAGAGAGAGAGAG 6447
QY 6014 AGTCTCTCTCTTATCAATGATGATTTGATGATGATGATGATGATGATGATGATGATGATG 6073
Db 6448 AGTGAAGTTCTTAAATGATGATTTGATGATGATGATGATGATGATGATGATGATGATG 6507
QY 6074 TGCCCTCTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6133
Db 6508 TGCCATTCCTTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6567
QY 6134 CTGATTTAATTA 6147
Db 6568 CTGATTTAATTA 6581

RESULT 4
US-09-880-107-3439
/ Sequence 3439, Application US/09880107
/ Patent No. US20020142981A1
/ GENERAL INFORMATION:
/ APPLICANT: Home, Darcie T.
/ APPLICANT: Vockley, Joseph G.
/ APPLICANT: Scheir, Uwe
/ APPLICANT: Gene Logic, Inc.
/ TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
/ FILE REFERENCE: 44921-5028-MO
/ CURRENT APPLICATION NUMBER: US/09/880,107
/ CURRENT FILING DATE: 2001-06-14
/ PRIOR APPLICATION NUMBER: US 60/211,379
/ PRIOR FILING DATE: 2000-06-14
/ PRIOR APPLICATION NUMBER: US 60/237,054
/ PRIOR FILING DATE: 2000-10-02
/ NUMBER OF SEQ ID NOS: 3950
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 3439
/ LENGTH: 6608
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: Genbank Accession No. US20020142981A1 U81607
US-09-880-107-3439

Query Match 34.8%; Score 2144.4; DB 10; Length 6608;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 4215; Conservative 0; Mismatches 1796; Indels 543; Gaps 45;

```

Db 522 GACTCTGAGATGTGAGGAAAGAGACTCCGATTAAGAGATGGCTACTAAGTCAAGCGTT 581
Qy 369 GTTGAAGATATCAACAAGGATGGGACGAGGAGACATCAGAAATAATTTGAACAGATCCCT 428
Db 582 GTTCACGACATCACAGATGATGGGACGAGGAGA---ACCGAATAATCGAACAGATTCCT 638
Qy 429 GCTTCAGAAAACATGTGGAAGAAATGGTACAGCCTGTGAGTCCAGGCTTAATGATGTT 488
Db 639 TCTTCAGAAAAGCAATTTAGAGAGCTAACAACACCCACTGAGTCCAGGCTTAATGATATT 698
Qy 489 GGCTTCAAGAAAGTATTTAAATTTGTTGTTTAAATTCACGGTGAAGAGATAAAAT 548
Db 699 GGAATTAAGAAAGGTGTTTAAGTTTGTGGCTTTAAATTCACGTGAAGAAAGGATAAGACA 758
Qy 549 GAAAGTCAGATCTCTCAACTACTCACTGTCAAGAGAGTGAAGGCGAAGGGGCGAGAA 608
Db 759 GAGAAGCCTGACACTGTCCAGCTACTCACTGTGAAGAAAGTGAAGGGGAGGAGCAGCA 818
Qy 609 GCCTCTGTCGGAGTCGAGACACAGAGCCAGTGTGAGACTGCGGTGCGAGAGTCA 668
Db 819 -----GGGGTGGCGACACAGAGACCCAGCCTTGGGGCT-----GGAGAAAGCA 863
Qy 669 GCATCCAAAGAAAGTGAGCTGAAGCAATCCACAGAGAGCAAGAGGACCCCTGAAGCAA 728
Db 864 GCATCCAAAGAAAGCGAAACCCAAACAATCTACAGAGAAACCCGAAGAGACCCCTGAAGCGT 923
Qy 729 GAACAGAGCAGCAGAGAAATCCCTCTCAAGCCGAATCTGATCAAGGGCGTGAAGAAAGAA 788
Db 924 GAGCAAGCCACGAGAAATTTCTCCCCACCCGNACTTGGCCAAAGCAGT---GGAGGAA 980
Qy 789 GCCAAAGATGAAGAGAGAAAGAAACAAGAGAAAGAGCCCAAGTCCCCAGATCCCCG 848
Db 981 TGCAAGAGGAGAGAGAGAGAAACAAGAAAGAAACCTAGCAAGTCTGCAGAAATCTCCG 1040
Qy 849 AGCAGCCAGTCAACAGTGAACAACATCTTCTTCAAGAGTTCTTCACTCAAGGTTGG 908
Db 1041 ACTAGTCCCGTGAACAGTGAACAGGATCAACCTTCAAAAAATTTCTTCACTCAAGGTTGG 1100
Qy 909 GCCGGCTGGCCAAAGAACAGCAGCTTCAAGAAATCAAAAGAGGATGATCTGGAATCTGCC 968
Db 1101 GCCGGCTGGCCNAAGAACCCAGTTTCAGGAAGCGGAAGAGATGAAGTGAAGCTTCA 1160
Qy 969 GAGAAGAAAGAGAGCAGAGCGAGAAAGTGAAGAGAGAGAAAGAAAGAAAGAACACAGAG 1028
Db 1161 GAGAAGAAAGAGAGAACAGAGCCAGAAAGTAGACACAGAAAGAGACGGAAGAGCAGAG 1220
Qy 1029 CCAGCCTCGAGGA-----GGAGGAGCCGCGAGAGAC 1061
Db 1221 GTTCCTCCGAGAAACTGACCGCTCCGAGCAAGCCACCCACAGGAGCCGCGAGAAAGT 1280
Qy 1062 ACAGACGAGGCGAGGTTGTCAAGCAGACTACGAGAAAGGTGAGCTGCCTTTTGAAGAGCCAG 1121
Db 1281 GCCACAGGCCCCGGTTATCAGCTGAATATAGAAAGTTGAGCTGCCCTCAGAGAGGCAA 1340
Qy 1122 GTTGGTGACTGAGGAGCATCTCAGAGAGAGTGTGCTCTCTTTGGCAACGGAAAGTGTTT 1181
Db 1341 GTCAAGTGCTCGCAGGAGACCTTCTGAAGAGAAACCTGTCTCGTTGGCGACAGAAAGTGT 1400
Qy 1182 GATCAGAGATGGAAGCCCAACAGAA---GTTGTTGAGAGGTTCCAGTGAACCCGTG 1238
Db 1401 GATCAGAAATAGAAAGTCCACNAAGAGAGTTGTGGCCGAAGTCCACGTGAGCACCCTG 1460
Qy 1239 GAGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGCTGAAGGGGCGTGGTGTAGAA 1298
Db 1461 GAGGAGAGACCGAGAGCAGAAA-----ACGAGGTGGAA 1496
Qy 1299 GGAACAGAGAAATCTTGGCCCCCTGAGAAACTGGTGAAGCCCGCAGGAGGTCCCCAGAGAA 1358
Db 1497 GAAACAGCAGGCTGTGTGCCAGCTGAAGAAATGGTTGGAATGGATGCAGAACCTCAGGAA 1556
Qy 1359 GCTGAGCCTGCTGAGGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGAGGAGAGACAC 1418

Db 1557 GCCGAACCTCCAAAGGAGCTGGTGAAGCTCAAAGAAACGTGTGTTTCCGAGAGGACCCCT 1616
Qy 1419 ACTCAACTGACAGACCTTAAGTCTCTGAAGAGAAGCGCTGCCCAAACCCAGAGAGGCATT 1478
Db 1617 ACAGAGGAGCTGACCTCAGTCTCTGATGAGAAGGTGCTGTCCAAACCCCCCGAAGGCGTT 1676
Qy 1479 GTCAAGTGAAGTGAAGTGTCTCTCTCAGAAAGAAATCAAGGTACAGGGAAGTCCCTTG 1538
Db 1677 GTGAGTGAAGTGAAGATGCTGTCTATCAAGGAGAGAAATGAAGGTGCGAGGAAGTCCACTA 1736
Qy 1539 AAGAAACTCTTTAGTAGTCTCAGGCTTAAAGAGCTGTCTGGAAAGAAAGCAGAGGGGAAA 1598
Db 1737 AAGAGCTTTTACCAGACTGGCTTAAAGAGCTTCTGAAAAGAAACAGAAAGGAAA 1796
Qy 1599 CGAGGAGGTGGGAGACGAGAGCCTTGAAGATACCAACACATTTCAACCGAATCCCCA 1658
Db 1797 AGAGGA---GGAGGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCGGATTTCTCCG 1853
Qy 1659 GAGAGTCTGATGAGCAGAAAGGAGAGAGCTCTGCGTGTCTCCCCGAGGAGCCTGAGGAG 1718
Db 1854 GACAGCCAGGAGGAGCAAAAAGGGCGAGAGCTCTGCTTCATCTCCCTGAGGAGCCCGAGGAG 1913
Qy 1719 ACCACGTGTCTGGAGAAAGGCGCTGGAAGACACCCAGATGGGAAAGCTTGAGGAAGGA 1778
Db 1914 ATCAAGTGTCTGAAAAGGCTTAGCCAGGTGCAGCAGATGGGAAAGCTGAAGAAGGA 1973
Qy 1779 ACTACTTCCGATGGAGAGAAAGAGAGAGGATCACTCCCTGGGCATCTTTCAAAAAG 1838
Db 1974 GCTACTTCCGATGGAGAGAAAGAGAGAGGTCTCACTCCCTGGGCATCATTTCAAAAAG 2033
Qy 1839 ATGTTGACACCCAAAGAAACGGGTCCGAAGACCTTCTGAGAGTGAAGAGGAGAAAGAGCTG 1898
Db 2034 ATGTTGACGCCCAGAAAGCGTTTAGACGCGCTTCGGAAGAGTGAATAAGAAAGATGAGCTG 2093
Qy 1899 GAGAGGTCAAGAGCGCCACTTGTCTCTCACTGATAGCACAGTGTGAGAAATCCAGAT 1958
Db 2094 GACAAGTCAAGAGCGCTACTTGTCTTCCACCGAGAGCAGCCTCTGAAATCAAGAA 2153
Qy 1959 GAAGTCAAAACTGTGTTGGTGAAGAACAAAAGCCAGAGAAACCAAGCGTAGGTTGGATACT 2018
Db 2154 GAATGAAGGGAGCGTGGAGAGCCAAAGCCGGAAGAACCAAGCGCAGGTGATAC 2213
Qy 2019 TCAGTGTCTTTGGAGACACTGATTTGTGTGTCGATCATCTCAAGAAAGAGAGCAAGAGGCA 2078
Db 2214 TCAGTATCTTTGGAGCTTTAAATTTGTGTGGATCATCTCAAGAAAGAGCAAGAGAGG 2273
Qy 2079 TCCTCTTCAGATGATGAAGAGGCGCCAGGACACTGGGAGGAGCAGTCAAGAGCAGAG 2138
Db 2274 TCCTCTTCTGATGAGGAAGGGGACCAAAAGCAATGGGAGAGACCAACCAAGAGCTGAT 2333
Qy 2139 GAGGCCAGCAAGAACAAAGCCGGAACAGAGCTGTTCTTCCAGACACCCAGGAGCAG 2198
Db 2334 GAGGCCGGAAGAACAAAGAGACGGGACAGACGGGATCTTGTGTTTCCAAAGACAT 2393
Qy 2199 GACCAAGCGCAAGAGTCTCTCAACCGAGCCAGCGGGAAGCCCTTCCGAAGGGGAAGGT 2258
Db 2394 GATCCAGGCGAGGAGTCTCTCCCGAGCAAGCTGGAAGCCCTACCGAAGGGGAGGCG 2453
Qy 2259 GTCTCCACTTTGGAGTCAATTTAAAGATTAGTCACTCCAGAAAGAAATCCAGTCAAAA 2318
Db 2454 GTTTCCACTTTGGAGTCAATTTAAAGGTTAGTCAAGCCAAAGAAAAATCAAGTCCAAG 2513
Qy 2319 CTGGAAGAGAAAGCCGAAGAC-----TCTAGTGTAGAGCAGTGTCCACTGAG 2366
Db 2514 CTGGAAGAGAAAGCGAAGACTCCATAGCTGGGTCTGTGTAGAACATTCCACTCCAGAC 2573
Qy 2367 ATCGAACCGAGTAGAGAAAGATCTTGGGTTTCCATTAAAGAAATTCATCCCCGAGCGCG 2426
Db 2574 ACTGAAACCCGGTAAAGAGAAATCTCTGGGTCTCAATCAAGAAAGTTTATTTCTTGGACGAAG 2633
Qy 2427 AAGAAAGGCGAGCGGGAAGAGCAAGACCACTGTGGAAGCTCAGGCGCCAGTGGAG 2486
Db 2634 AAGAAAGGCGCAGATGGGAAACCAAGAACAGGCCCTCTGTTGAAGACGAGGCGCCACAGGG 2693

QY 2487 ATAAATAGAGCAGACCTTAATGTCCAGCCGTCGTCTGTCTGAGTATTAATGCACTG 2546
Db 2694 GCCAACCAAGATGACTGTGATGTCCCGGCGGTGTCTGTCTGATGATGATGTCTGTA 2753
QY 2547 GAGAGGAGAAAGATGAG-----AAGCCAGGGGAAATAGGGAGCTGCCAGCTGTG 2597
Db 2754 GAAAGGGAGAAATGAGAGCAGACAGACCCAAAAAGGCGCAGAGACAGCCGAGCAGAG 2813
QY 2598 GGGGCTGTGATCTGTCCGAGAGCTCAGTAAGACTGTGTCCAACCTGTGATGTGCGCA 2657
Db 2814 GCAGCCACTGAGGTGTCCAGAGAGCTCAGCGAGGTCAAGTTTCAATGATGAGCAGACCT 2873
QY 2658 GTCAATTGATGGAGCCAGGGCAGTCACTCAGTCCAAAGCCGCTCTCTTGGATATCC 2717
Db 2874 GTCCCTGACGGAGCGAGGGCAGCTATCTATTGAAAGAAAGTCTCTCTTGTGGATATCT 2933
QY 2718 GCTTCCGTAACAGAACCTCTGAAACACACAGCGGGAGAAAGCATGCGCACTGTGAAAG 2777
Db 2934 GCTTCAGTGAAGAACCTCTTGAACTAGTAAGCTGAAGCCGCACTGTTAACAGAGAG 2993
QY 2778 GTCACTGAAAAAGACATCATTTGC---AGAGAAACTCTGTGTCTCACCCAGACGTTTACA 2834
Db 2994 GTATTGGAAAAGATGATTTGCAAGAGAGAACCCCCAGGTTACTGAACTGTGCA 3053
QY 2835 GAGGCTAAAGATGCCATGACGACATGCTCACCACTGAATGATGATTTCACTCAGAACCT 2894
Db 3054 GAGAAACAGAGAGCGCGGGGCGACACGGTCTGTATTGAGGGCGAAATTCACCCCCMACT 3113
QY 2895 GTGACAGCCACAGAGACCTCAGAGGCTCTCCGTACTGAAGAAATTACGAGACATCGGG 2954
Db 3114 GTGACAGTGAAGAACTGACAGGGCATTTGGGTTCGAAAGAGAAACCGAAMACCTGTCT 3173
QY 2955 GCCCAAGAGACCAACAGACATGCTGTCCGAGTTTCCAGCTGACTGACTCTCCACAGACA 3014
Db 3174 GCTGAAGAGACCAAGAAATGCTGTCAAGAGTCTCCAGTTTAAACGACTCCCAACAC 3233
QY 3015 ACAAGAGAACCCCACTTCCAGAGGTAGAGAGTGTGTCTGATACAGAAAGAG 3074
Db 3234 ACAGAGAGGCGCACTCCGGTGCAGAGGTGAAAGGTGCTGATCATGAAATGAAAGCA 3293
QY 3075 GAGCGCCAGACGAGCCATCTCCAAGCCGTTGACAGCAAGGTTAAAGAGAGTCCAG 3134
Db 3294 GAGAGGCGGACTCAAGAGTCTTCAAGCAGTGGCAAGAAAGTAAAGAGAAATCCAG 3353
QY 3135 GTGCTGTCAACCC---CAGACTGTGCAGAGAAAGGGGTCAAAAGCACTGAGAAAGTTGAG 3191
Db 3354 CTGCTGTGCACCGGTGGGCGAGAAATGTGCTTCAAGCTGTGCAAGAGCAGAGCAGAA 3413
QY 3192 GAGGTAGAGAGACTCCGAAGTGTGCTTCCGAGAAAGAGAGAGAGCTTATGCCGAA 3251
Db 3414 AGACCAAGAGAGCAGGCTGAGAGCTGTGGGCTGGAAGAAAGAACCGAATGTATGTTGAAA 3473
QY 3252 GAGCCCGTGCAGAGAGCTGAGAGTGAATCTTGACAGGGCTGTGAGATCGAAGGCT 3311
Db 3474 GTAGATGCTCAGAGAGCAAAATCTAAGCTTTTACAAAGGAGTGTGGGCGACACC 3533
QY 3312 ACTCCAGAGAGCTTG---AAGTTCTGAAGTCAAGGCGAGATGAGAC-----CAT 3359
Db 3534 ACCCAGAAAGCTTTGAAAGAGCTCTCAAGTCAAGAGAGATGAGTCCAGTAGCTT 3593
QY 3360 GTCCGCAAGTGCA-----GTTTATCAAGCTTCCAGCAGCTGATGANA 3401
Db 3594 GTAAACCATTTGTCAAGCCGAAACTTAACTGTGGGTAAATATCAAGAGATGTGTGTA 3653
QY 3402 CAGGCGGTGGCCCTGAGTCAATCCGAAACTTGAACAGACATGAGACAAATGAGAGCT 3461
Db 3654 CAGGCTATCCCTGACTGTGGTGAAGCCCTTACAGAGAGTGAAGCTGATGAGAGCAC 3713
QY 3462 CCCTTAGAGATTTCAGACTGATGAGGACACAGCAAGATGAAACCATTTGACAGCCAG 3521
Db 3714 CCGGTAGCGACTTTGACGACACAGGACACACCCAGAAAGAGAGATTGTGAAATCCAT 3773

QY 3522 GACAGTAAAGCACTGACAGTGTGAGGAGTCAAGGTCACAGAGTCAAGAAAGAGAGCGGCTACT 3581
Db 3774 GAGAGAAATGAGTGTGCAATTTGTAATCCAGTCAAGGGGCAACAGAGACAGAGCGATTCT 3833
QY 3582 GCTCAGAAAGAGAGACCTTGACACTCTAATTAATGTTCCAGGCCAGAGAAACATGGG 3641
Db 3834 GCACAGAAAGAGAGGCTCCAGAGACCTTCCAGTTTTGTGTTCCAGAAAGAAATCAAAAGAA 3893
QY 3642 GAAGAACCAAGAGAGATGTTCTTGAACCTTACACAGCAAGACTTACTGCTGCAGCCGTG 3701
Db 3894 CAATCAAAAGATGAAAGACCTTGAACATACAGATTAAGAGGGGTGATGAGTGAATCTTA 3953
QY 3702 CCGTGTGGCAAAAGACTGAGGTGGTCAAGAGGCTGA-----GTTTACTGTGTTG 3752
Db 3954 TTCATTTGTTCAAAGACTGAGGGGACTCAAGAGCTGACAGTATGCTGATGAGAAAAAC 4013
QY 3753 GATGAGAAAAAGTCAAAAGAAACAGAGAGTGTGTTGACACTGTG-----ACCCAAC 3806
Db 4014 AAAGACGTACATTTTTCGAAAGACTTGAAGGCTGTATGACACAGGCATTAACAGTCAGT 4073
QY 3807 AGTCAAAAGCTGCTGATGTGACATATGACATGATGATGATGAGTGGCCGGGTGTGAG 3866
Db 4074 CGGAAAAAGTCACTGAAGTGGCCCTTAAAGGTGAAGGACAGAGAAAGCTGAATGTAAA 4133
QY 3867 GAAAGGAGAGTACTGAAGTGCAGAG-----TCTTAGCCTGAGAGAGGA 3911
Db 4134 AAGATGATGCTCTTGAACCTGACAGATCAGCTTAAGTCTCTTCAATCCCGTGAAGAGA 4193
QY 3912 GAGATGAAACTGACTGTGAAAGAGAGAAAGGAGACAAAGCCAGAGCAAGTGAAGTGA 3971
Db 4194 GAGATGTGATTTCAAGTGCAGAAAGGAGAGAAACAGAACAGAGCCACCATGTGAATGAA 4253
QY 3972 GAAGG---TGAGCAGAGAAACAGCCGCTCTGAGCATGAAGAACCTTACGGAGCCAGTC 4028
Db 4254 GAGAAAGCTTGAAGCAGAAACGCTGTTAACCTATCTGAAGAGTCAATGAACACTCTCTC 4313
QY 4029 CTGACACTTGCATGCCAGCTCAGAGAGGGGAGAGGACTGGGAAGCTTGAAGAGAGC 4088
Db 4314 CAGACAGTGAATGTGCCATATATGATGGGCAAGAAAGTCAACAGTGTGAAAGAGAGC 4373
QY 4089 CTTTCTC---TCCAGACCAAGCAAAAGCAGGTTGCAATGAGTTTCAAGTTCAAAAGCTGTG 4145
Db 4374 CTTCTCCCTGCTTAAGTCAAGAGAGAGGAGATGACACAAATTTCAAGTTCAAGGCTCT 4433
QY 4146 GACACACAGTCACTCAACACAGCAGAACTGTGAGAAA-----GTCATAGAAACGGTT 4199
Db 4434 GAGGCATATTCACCTTAAACGCGCTGCAGAGAGGAAAGTCTTATGAGGAAATGACC 4493
QY 4280 GTGATTTCAAGAGACAGGTGAAGAGTCCAGAGTGTGATGGTGCACATTATTCACAGCTGAG 4259
Db 4494 AACATTTTGAAGACAGGTGAAGCTTGAAGCTTCAAGGTGACATTTTATGTTCTGGAAGAG 4553
QY 4260 AAGTCTTTCGCAAGGGTGGCACTGTGACTTTCAGCATGACAGAGACACACGGTACCCCTG 4319
Db 4554 AATCTCTGAAAAAAATGAAAGCTTTGGCCGCTCATCCAGGGGAAAGTGTGTGCCACA 4613
QY 4320 GGGCTGAGTCTCAGGCGAATTCATCCATCATATGTAATCTCTGCTCTGAAAAGCAC 4379
Db 4614 GGGCCCGACTGTCAAGGCAAAATTCAGACCAAGTGAATGATATCTGTACTTACCAAGAAAGGC 4673
QY 4380 CTACATCTTGACTTACAGAGGAAATTAAGGCATCCAGAGAGAGCGATTCAGAGGAAG 4439
Db 4674 TTAAGTTCGACCTGGAAGGAGAGAAACCAATTCATGAAAGTGAAGTCAAGATGAATTC 4733
QY 4440 GACAAGCCAGATGTGTGCTGATGCTGACGGAAGAGAGATACAGCAATGCAAAAAATGC 4499
Db 4734 GATGAGCAGGTTGTCTGTGC-----AGAGGTCAAAAGTGAATGATGACATTTG---AGGAT 4784
QY 4500 CTCAAGGCTGAACCTGAGATCTTGAACCTTGAAGATTAAGACCAACAGATTGTGCTGAAC 4559
Db 4785 TTAGAGCTTGAATTTGGATTTTGAACCTTGAAGCAAAACAGATTAACCTTCCAAAAAC 4844
QY 4560 GTCAATTACAGACGCGTGTGACCAAGTTCGACAGTAC---AGAAACAGCCCGGAAACCTAT 4616

; SEQ ID NO 2
; LENGTH: 5346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-877-2

Query Match 31.6%; Score 1945.6; DB 10; Length 5346;
Best Local Similarity 67.5%; Pred. No. 0;
Matches 3253; Conservative 0; Mismatches 1324; Indels 243; Gaps 27;

QY 33 ATGGGCGGAGGAGTTCACCGAGGAGGAGCCCGAGCGAGCCGGC---GGGAGCGAC 89
Db 1 ATGGGCGGAGGAGTTCACCGAGGAGGAGCCCGAGCGAGCCCGAGCGAGCTCC 60
QY 90 ACGCGGAGGAGGAGTTCAGTGGCCATGGGCCCGAGCTGAAAGCTCGGA----- 143
Db 61 ACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 144 ---GAGCTGGAGAGCCCGCGGAGCGCGGAGCCCGGAGCTCCAGAGAAATGGC 200
Db 121 ACCGGGAGAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 201 CAGCTGTCTTCGTCAGCGGAGTGTGAACAAGAGATGTCTATGTCCAAAGAGAAAC 260
Db 181 CAGCTGTCCATCATATGGCTTACTAGCAAGATGAGCTCAAGAGGAGTAC 240
QY 261 CAGGAGG-----GCAGAGGAA 278
Db 241 CTAAATGCGCAGAAAGAGAGCCCTGAACGGTCAAGAGCCCTTAAACAGCAGAGAGAA 300
QY 279 GAGTGTCTTGATGAGAGATGTTGGAACAAGAGATGTAAGAGAAACCA 338
Db 301 GAAAGCATTTGTCAACGAGGATGGAGAGAGACTGTGAAGATGTAAGAGAAACCTCC 360
QY 339 GTTGAAGAAATGGCGGCAACTCCAAGCTGTTGAAGATATCAAAAGATGGGAGAG 398
Db 361 GATTAAGAGATGGCTACTTAAGTCAAGCGGTTTTCAGACATACAGATATATGGAGAG 420
QY 399 GAGCATCAGAAATATTAAGACAGATCCCTGCTTCAAGAAACATGTGAGAAATGTA 458
Db 421 GAGA---ACCGAAATATGGAACAGATTCCTTCTCAAGAAACATTTGAAGAGTAAACA 477
QY 459 CAGCTGTGAGTCCCGAGCTATATGTTGGCTTCAAGAAAGTATTTATTTTGGT 518
Db 478 CAACCCACTGAGTCCAGGCTATATGATTTGATTTAAGAGTGTATTAAGTTTGTGCG 537
QY 519 TTTAATTCACGCTGAAGAGATTAATAATGAAGTCAAGTACTGTCCAACTACTCAT 578
Db 538 TTTAATTCACGCTGAAGAGATTAAGAGAGAGAGCTGACATGTCTCAAGCTTACTACT 597
QY 579 GTCAAGAGATGAAGGCGCAAGGCGAGAGGCTGTGTGAGCTGAGAGCCAGAGAG 638
Db 598 GTGAAGAAAGATGAAGGAGGAGGAGAGCA-----GGGGCTGGGCAACCAAGAGAG 648
QY 639 CCCAGTGTGAGAGCTCCGTCGGAAGTACAGATCCAAAGAAATGAGCTGAAGCAATCC 698
Db 649 CCCAGCTTGGGGCT-----GGAAGAGCAGCATCCAAAGAAAGCGAACCCAAACATCT 702
QY 699 ACAGAGAGAGAGAGGAGCCCTGAAGCAAGAAACAGAGAGACAGAAATCCCCCTTCA 758
Db 703 ACAG 762
QY 759 GCGGATCTGATCAAGCGGCTGAGAGAGAGCAAAAGATGAAGAGAGAGAGAGAGAG 818
Db 763 GCGGATCTGGGCA---GCAGTGAGAGATCAAAAGAGAGAGAGAGAGAGAGAGAGAG 819
QY 819 AAAGAGCCCAAGAGTCCCGAGAGATCCCGAGAGCCAGTCAACAGTGAAGCAATCT 878
Db 820 AAAGAGCCCAAGAGTCTGCAAGATCTCCGACTAGTCCGTCAGCCAGTGAAGAGAGAT 879
QY 879 TCCTTCAGAGATTTCTTCACTCAGCGTTGGGCGGCTGGCGAGAGAGAGAGAGAGAG 938
Db 880 ACCTTCAGAGATTTCTTCACTCAGAGGTGGGCGGCTGGCGAGAGAGAGAGAGAGAG 939

QY 939 AAATCAAAAGAGATGATCTCGAAACTCCCGAGAGAGAGAGAGAGAGAGAGAGAG 998
Db 940 AAAGCGAAGAGATGAAGTGAAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999
QY 999 GTAGACGAG 1042
Db 1000 GTAGACACAG 1059
QY 1043 -----GCAGAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1091
Db 1060 CAAGCCACCCAG 1119
QY 1092 GAGAAAGTGAAGTCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1151
Db 1120 GAGAAAGTGAAGTCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1179
QY 1152 AAGTGTCTCTTGAAG 1208
Db 1180 AAACCTGCTCTTGAAG 1239
QY 1209 GTTGTGAGAGAGTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1268
Db 1240 GTTGTGAGAGAGTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1295
QY 1269 GGAAGGCTGAAGGGGCGTGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1328
Db 1296 -----GAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1335
QY 1329 CTGGCTGAGCCCGAGAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1388
Db 1336 TTGTGTGAATGATTCAGAACTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1395
QY 1389 AGAGAGATGTGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1448
Db 1396 AAAGAAAGTGTGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1455
QY 1449 AAGAGCTGCCCAAAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1508
Db 1456 AAGGTGTGTCAAAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1515
QY 1509 GAAAGATCAAGGTACAG 1568
Db 1516 GAGAGATGAAGGTGAG 1575
QY 1569 AAGGTGTGAG 1628
Db 1576 AAGCTTCTGAG 1632
QY 1629 GAATTCACACATTCACACAGAGATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1688
Db 1633 GAGCAGACTCAGGTTCAG 1692
QY 1689 TCTGTGTGTCTCCCGAG 1748
Db 1693 TCTGTGTGTCTCCCGAG 1752
QY 1749 GCACCCAG 1808
Db 1753 GTGACAG 1812
QY 1809 GGGATCACTCCCTGGGAGTCTCTTCAAAAGATGTGAGAGAGAGAGAGAGAGAGAGAG 1868
Db 1813 GGTGTCACTCCCTGGGAGTCTCTTCAAAAGATGTGAGAGAGAGAGAGAGAGAGAGAG 1872
QY 1869 CTTCTGAG 1928
Db 1873 CTTCTGAG 1932
QY 1929 ACTGTAGCAGAGTGTGAG 1988
Db 1933 ACCGAG 1992

QY 1989 CCAGGNAACAAAGCGTAGGTGGATACCTTCAGTGTCTGGGAAGCACTGATTTGTGTC 2048
Db 1993 CCGGAAGAACAAAGCGAAGGTGGATACCTTCAGTATCTTTGGGAAGCTTTAAATTTGTGTG 2052
QY 2049 GGATCATCCAAAGAGAGACAAAGAGGATCCCTCTTCAGATGATGAAGGAGGCCCAAGG 2108
Db 2053 GGATCATCCAAAGAAAGACAAAGAGGATCCCTCTTCAGTATGAGGAAGGGGACCAAAA 2112
QY 2109 AACTGGGAGGGGACAGTTCACAGACGACGAGGAGGCCCAAGACAAAGAAAGCCGGAACA 2168
Db 2113 GCAATGGGAGGAGACCAACAGAAAGCTGATGAGGCCGGAAGAAAGACAAAGAGAGCGGAC 2172
QY 2169 GACGCTCTTCCTCCAGCAGCCAGGAGCAGACCAAGGCCAGGAAGTTCCTCACCAGG 2228
Db 2173 GACGGGATCTCTTGTGTCTCCAAAGACATGATCCAGGGCAGGGAAGTTCCTCCCGGAG 2232
QY 2229 CCAGCGGAAGCCCTTCGGAAGGGGAAGGTGTCTCCTCTGGGAGTCAATTTAAAGATTA 2288
Db 2233 CAGCTGGAGCCCTACCGAAGGGGAGGGGCTTTCCACCTGGGAGTCAATTTAAAGGTTA 2292
QY 2289 GTCACTCCAAAGAAAAATCCAAAGTCAAAACTGGAAGAGAAAGCCGAAGAC- 2339
Db 2293 GTACGCGCAAGAAAAATCAAAGTCCAAGCTGGAAGAGAAAAAGCGAAGACTCCATAGCT 2352
QY 2340 ---TCTAGTGTAGAGCAGTTGTCCACTGAGATCGAAACGAGTAGAGAAATCTTGGGTT 2396
Db 2353 GGGTCTGGTGTAGAACATTCACCTCCAGACACTGAACCCCGGTAAAGAAATCTCGGGTC 2412
QY 2397 TCCATTAAGAAATTCATCCCGGAGCGCGGAAGAAAGGGCAGACGGGAGCAAGAACAA 2456
Db 2413 TCAATCAAGAAAGTTTATCTCTGGACGAGGAGAAAGGCGCAGATGGGAACAAGAACAA 2472
QY 2457 GCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGACCCCTAATGTCCAGCC 2516
Db 2473 GCCCTGTGGAAGAGCAGGGCCAAACAGGGGCCAAACGAAGTACTCTGATGTCCCGGCC 2532
QY 2517 GTGTGCTCTGTCTGAGTATATGATGAGTGTGAGAGGAGGAAGATGG- 2567
Db 2533 GTGTGCTCTGTCTGAGTATGATGCTGTGAAAGGGGAGAAATGGAGGCGACAGCAAGCC 2592
QY 2568 CAGGGGAATACGGAGTGTCCCGAGCTGTGGGGCTGTGTACGTGTCCGAGGAGCTCAGT 2627
Db 2593 CAAAAGCGCAGAGCAGCCCGACAGAGCGACGCACTGAGGTGTCCAAGGAGCTCAGC 2652
QY 2628 AAGACTGTGTCCACTGTGAGTGTGCGCAGTCAATGATGGGACAGGGCAGTCAACAGT 2687
Db 2653 GAGAGTCAGGTTTCATATGATGCGCAGCAGCTGTCTGTCGCGGACAGGCGCAGCTACCAT 2712
QY 2688 GTCGAAGAGCGGTCTCCTTGTGTGATATCCGCTTCCGTTAAGCAAGACCTCTTGAACACACA 2747
Db 2713 ATTGAAGAAAGGTCTCCTTCTTGATATCTGCTTTCAGTGACAGAACCTCTTTGAACAAGTA 2772
QY 2748 GCGGGAGAACGATGCGACCTGTGTAAGAGGTCACTGAAAGACATCATTCG- 2804
Db 2773 GAAGCTGAAGCCGACCTGTGTAAGTGTGTAAGAGAGAGTGTGTAAGAGAGAA 2832
QY 2805 GAAACTCTGTGTCCAGCAGACCTTACAGAGGGTAAAGATGCCATGACGACATCGTC 2864
Db 2833 GAACCCCGACGGTTACTGAACCTCTGCCAGAGAACAGAGNGGCCCGGGGGACACCGTC 2892
QY 2865 ACCAGTGAAGTGAATTTCACTCAGAGAGCTGTGACGCCACAGAGACCTCAGAGGCTCTC 2924
Db 2893 GTTAGTGAGGCGGAATTTGACCCCGAAGCTGTGACAGCTGCAGAAAATCGCAGGSCCAT 2952
QY 2925 CGTACTGAAGAGTTACGGAAGCATCGGGGCCCAAGAGACCAAGACATGTTGTCGCA 2984
Db 2953 GGTTCGAAGAGGAACCGAAGCATCTGCTGTGAAGAGACCAAGAGAAATGGTGTACGA 3012
QY 2985 GTTTCCAGCTGACTGCCAGACACCAAGAGGAAGCCAGCCAGTTTCAGGAGGTA 3044
Db 3013 GTCTCCAGTTAACCGACTCCCGACAGACCAACAGAGGAGGCCACTCCGGTGCAGAGGTG 3072
QY 3045 GAGAGTGTGTGTAGATACAGAAAGAGAGGAGCGCCAGAGCGAGGCCATCTCTCAAGCC 3104

Db 3073 GAAGGTGGCGGTACTGACATAGAAAGCAAGAGAGCGGACTCAAGAGGTCTCTCCAGGCA 3132
QY 3105 GTTTCAGACAAAGGTGAAGAGGAGTCCAGGTGCTCTCAACC---CAGACTGTGCAGAGA 3161
Db 3133 GTGCAGAAAAGTGAAGAGGAAATCCAGCTGCTGGCACCGGTGGCCAGAAGTGTG 3192
QY 3162 ACGGGTCAAAAGCACATGGAGAAAGTTGAGGAGGTAGAGGAGACTCCGAAGTGTGGCT 3221
Db 3193 CTTTCAGCTGTGCAGAGCAGAGGCGAGAAAGACCAAGAGCAGGCTGAAGCGTCGGGT 3252
QY 3222 TCGGAGAAAGAGAGGACGTTATGCCGAAGAGCCCGTGCAGNAAGCTGCAGCTGAGCAT 3281
Db 3253 CTGAAGAAAGAGACGATGTAGTGTGAAAGTAGATGCTCAGGAGGCAAAACCTGAGCCT 3312
QY 3282 CTTTCACAGGGCTCTGAGACTGGACAGCTGACAGGCTACTCCAGAGAGCCTTG---AAGTTCCTGAA 3338
Db 3313 TTTTACACAGGGAAGGTGGTGGGCGACACCCCGAGAAAGCTTTGAAAAGCTCTCTCAA 3372
QY 3339 GTCACGGACAGATGTA-----GACCATGTGCGCCACGTGCCA----- 3373
Db 3373 GTCACAGAGAGCATAGAGTCCAGTGAGCTTGTAAACCACTTGTCAAGCGCGAAACCTTAGCT 3432
QY 3374 --GGTTATCAAGCTCCAGCAGCTGATGGAAACAGCCCGTGGCCCTGAGTCTATCGGAAC 3431
Db 3433 GGGGTAAAATPCACAGGAGATGGTGTATGGAAACAGGCTATCCCCCTCTGACTCGGTGGAAAC 3492
QY 3432 TTGACAGACAGTGCAGAAATGGAAGCACTCCCTTAGCAGATTCAGACACTGCAGATGGG 3491
Db 3493 CTTACAGACAGTGAAGACTGATGGAAGCACCCCGTAGCCGACTTTGACGCCACAGGCACA 3552
QY 3492 ACACAGCAAGATGAAACCATTTGACAGCAGGACAGTAAAGCCACTGCAGCTGTCCAGCAG 3551
Db 3553 ACCAGAAAGACGAGATTGTGGAATCCATGAGGAGAAATGAGGTGCGCATCTGTTACCCAG 3612
QY 3552 TCACAGTTCACAGAAAGAGGCGGTACTGTCTCAGAAAGAGAGCTTCGACACTACCT 3611
Db 3613 TCAGGGGGCACAGAAAGCAGAGGAGTTCCTGTCACAGAAAGAGAGGCTTCACGACCTTCC 3672
QY 3612 AATAATGTTCCAGCCAGGAAGAAACATGGGGAAGAACCAAGAAAGAGATGTTCTTTGAACCT 3671
Db 3673 AGTTTGTGTTCAGGAGAAACTAAGAAACAATCAAGATGGAAGACACTCTAGACAT 3732
QY 3672 ACACAGCAAGAGCTTACTGTGCAGCGCTGCCCGTTCTGGCAAGAGCTGAGGTGGGTCAA 3731
Db 3733 ACAGATAAGAGAGTGTCAAGTGAACCTGTATCCATCTGTCAAAGACTGAGGGGACTCAA 3792
QY 3732 GAGGGTGA-----GGTTGACTGGTGGATGGAGAAAGTCAAGAAAGAACAGGAG 3782
Db 3793 GAGGCTGACCATGTGTGATGAGAAACCAAGAGCGTACCATTTTTCGAAGGACTTTGAG 3852
QY 3783 GTGTTGTACACTCTGG-----ACCCACAGTCAAAAGGCTGTGATGTGACATATGAC 3836
Db 3853 GGGTCTATAGACACAGCATACAGTCAAGTCCGGGAAAGGTCACTGAAGTTGCCCTTAA 3912
QY 3837 AGTCAAGTGTGAGTGGCGGTGTTCAGAAAAGAGAGTACTGAAAGTGCAGAG---- 3892
Db 3913 GGTGAAGGGACAGAAAGCTGATGTAAAAGGATGATGCTCTTGAACCTGCAGAGTCA 3972
QY 3893 -----TCTTACCTTGGAGGGGAGAGATGGAACCTGACGTTGAAAAAGGAGAA 3941
Db 3973 GCTAAGTCTCTCCATCCCTCCGCTGGAGAGAGATGGTAGTTCAAAGTCCGAAGAGGAGAA 4032
QY 3942 AGGAGACAAAGCCAGAGCAAGTGAAGTGAAG---AAGGTGAGCAGGAAACAGCGCTCCT 3998
Db 4033 ACAGAAAGCAGAGCCCAACCCATGTGAATGAAGAGAACTTTGAGACCGAAACAGCTGTACC 4092
QY 3999 GAGCATGAAGAAACCTACGGGAAGCCAGTCTGACACTTTCACATGACATGCCAGCTCAGAGAG 4058
Db 4093 GTATCTGAAGAGTCAAGTAGCAGCTCCTCCAGACAGTGAATGTGCCCATCATAGATGG 4152
QY 4059 GGAAGGCACTGGGAAGCCCTTGGAGGAAGCCCTTCTC---TCCCAGACCAAGACAAAGCA 4115


```

Db 4153 GCAAGAGATGACAGCTTTGGAAGAGCCCTCTCTCTGCTTGTGTAAGAGAGAGGCA 4212
Qy 4116 GGTTCATAGAGGTTCAAGTTCAAGCCTGACACACAGTCACTCAAAACGACAGAGCT 4175
Db 4213 GTATGACCAAAATTCAGATTCAAGGCTCTGAGGATCATTCATCTAAACGGGCTGCA 4272
Qy 4176 GTGAAAAAGGTCAT-----AGAAACGGTTGATTTCAAGACAGAGTGAAGTCCAGAG 4229
Db 4273 GAGGAGGAAAAAGCTTTGAGGAAAACTGCCAACATTTTGAAGAAACGTTGAGAG 4332
Qy 4270 TGTGAGGAGACACTTATTAACAGCTGAGAAAGTCTCTGCAACGGGTGGCCACTGACT 4289
Db 4333 CCTGAGGTTGACATTTAGTTTGTGAAAGAAATCTCTGAAAAAAATGAAGACTTTGCC 4392
Qy 4290 CTTTCAGATGACAGAGACACGCTACCCCTGGGCTGAGTCTCAGGCAAGATCCATCCCA 4349
Db 4333 GCTCATTCAGGGGAGAGATGCTGTGCCACAGGGCCGACTGTCAGGCAAAATCGACACCA 4452
Qy 4350 ATCATATGTAATCTCTGCTCTCTGAAAGCACCCTTACATCTGACCTTACAGAGAAATTAAGC 4409
Db 4453 GTGATAGTATCTGCTTACTTACCAAGAAAGGCTTAAAGTTCCGACCTGGAAGAGAAAAAC 4512
Qy 4410 GCATCCCAAGAGAGAGCATCAAGAGAAAGACCAAGCCAGATGCTGCTGATGCTGAC 4469
Db 4513 ACATCTGAAAGTGAAGTCAAGATGAATGATGATGACAGGTTGCTTGC-----AGAG 4566
Qy 4470 GGCAAGAGAGATGACAGCAATCGAAAAAGTCTCAAGGCTGAACCTGAGATCTGGAATTT 4529
Db 4567 GTCAAGATGAGTGTGACATTTG---AGATTTAGAGCCTGAAAAATGGGATTTTGGAACTT 4623
Qy 4530 GAGAGTAAGACACAGATTTGTCTGAAAGCTTATTCAGACAGCCGCTTACCAAGTTGCA 4589
Db 4624 GAGACCAAAAGCAGTAACCTGTCCAAAAATCATTCACACAGCCGTTTACCAAGTTTGA 4683
Qy 4590 CGTAC---AGAAACAGCCCGGAAACTCATGCTTATGATTCACAGCCAGGTTCTGTGA 4646
Db 4684 CGTACGAAAGAAACAGCCACCCGAAATGTTGACGTCTGATTTACAGACACAGCTACGTTG 4743
Qy 4647 TGCAGGCTTGACAGCAGGGA 4666
Db 4744 ATAAAGCTGACAGCCAGGA 4763

```

RESULT 6
US-09-917-800A-1223/C

; Sequence 1223, Application US/09917800A
; Patent No. US20020119462A1

; GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna

; APPLICANT: Porter, Mark

; APPLICANT: Johnson, Kory

; APPLICANT: Castle, Arthur

; APPLICANT: Elashoff, Michael

; TITLE OF INVENTION: Molecular Toxicology Modeling

; FILE REFERENCE: 44921-5038-US

; CURRENT APPLICATION NUMBER: US/09/917,800A

; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/222,040

; PRIOR FILING DATE: 2000-07-31

; PRIOR APPLICATION NUMBER: US 60/222,880

; PRIOR FILING DATE: 2000-11-02

; PRIOR APPLICATION NUMBER: US 60/290,029

; PRIOR FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: US 60/290,645

; PRIOR FILING DATE: 2001-05-15

; PRIOR APPLICATION NUMBER: US 60/292,336

; PRIOR FILING DATE: 2001-05-22

; PRIOR APPLICATION NUMBER: US 60/295,798

; PRIOR FILING DATE: 2001-06-06

; PRIOR APPLICATION NUMBER: US 60/297,457

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,884

```

; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1223
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 A123818
US-09-917-800A-1223

```

Query Match 8.9%; Score 550; DB 10; Length 563;
Best Local Similarity 99.1%; Pred. No. 1.9e-135; Indels 0; Gaps 0;
Matches 553; Conservative 0; Mismatches 5;

```

Qy 5601 ATAGCAGCGCTTTGAAAAAGCTCCAGGCTCACTGTAACCTGACGAAACAGATACATT 5660
Db 563 AAAGGAGCGCTTTGAAAAAGCTCCAGGCTCACTGTAACCTGACGAAACAGATACATT 504
Qy 5661 CTTGGCAAGAGAGACAAAGTCTTTTAAAGTTTACTGATGCTTAAATCTGTGGCTTCT 5720
Db 503 CTTGGCAAGAGAGACAAAGTCTTTTAAAGTTTACTGATGCTTAAATCTGTGGCTTCT 444
Qy 5721 AGTCTCTGAAAGTGTGTTTCTCTATGACACAGGAGCTGAGAAATPAAACCCCATTT 5780
Db 443 AGTCTCTGAAAGTGTGTTTCTCTATGACACAGGAGCTGAGAAATPAAACCCCATTT 384
Qy 5781 TGAACATCCAGATGATGCCAATTAATACATGATTTTTCCTTAAAGCCCTGTCTTAAATGATGGA 5840
Db 383 TGAACATCCAGATGATGCCAATTAATACATGATTTTTCCTTAAAGCCCTGTCTTAAATGATGGA 324
Qy 5841 TCCAGTTGAAAGAGAGTCTCTGTGTGTCAGATTAAGCCCTGTCTTAAATGATGGA 5900
Db 323 TCCAGTTGAAAGAGAGTCTCTGTGTGTCAGATTAAGCCCTGTCTTAAATGATGGA 264
Qy 5901 CAATGATGTCCTTAAGCCATGAGATGTTCTTAAAGCAGAAAGAAATCTGTGTGATG 5960
Db 263 CAATGATGTCCTTAAGCCATGAGATGTTCTTAAAGCAGAAAGAAATCTGTGTGATG 204
Qy 5961 TTTTGTATGATCTCTCTATGCTGACCGAATTCATATGACATGCAAGTGAAGTCTGTG 6020
Db 203 TTTTGTATGATCTCTCTATGCTGACCGAATTCATATGACATGCAAGTGAAGTCTGTG 144
Qy 6021 TTTTGTATGATCTCTCTATGCTGACCGAATTCATATGACATGCAAGTGAAGTCTGTG 6080
Db 143 TTTTGTATGATCTCTCTATGCTGACCGAATTCATATGACATGCAAGTGAAGTCTGTG 84
Qy 6081 TCTTTAAGACAAATGTCATTAATGATGCTGCTTGGATTAATGATTTGACAACTGATTT 6140
Db 83 TCTTTAAGACAAATGTCATTAATGATGCTGCTTGGATTAATGATTTGACAACTGATTT 24
Qy 6141 AAATAACATATTTGACT 6158
Db 23 AAATAACATATTTGCTT 6

```

RESULT 7
US-09-917-800A-1475

; Sequence 1475, Application US/09917800A

; Patent No. US20020119462A1

; GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna

; APPLICANT: Porter, Mark

; APPLICANT: Johnson, Kory

; APPLICANT: Castle, Arthur

; APPLICANT: Elashoff, Michael

; TITLE OF INVENTION: Molecular Toxicology Modeling

; FILE REFERENCE: 44921-5038-US

; CURRENT APPLICATION NUMBER: US/09/917,800A

; CURRENT FILING DATE: 2001-07-31

RESULT 14
 US-09-864-761-20733
 Sequence 20733, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aeomica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30

Query Match	1.2#	Score	71	DB	10	Length	575
Best Local Similarity	53.3#	Pred.	No.	1e-08			
Matches	172	Conservative	0	Mismatches	150	Indels	1
							Gaps
							1
QY	969	GAGAAAGAAAAGAGCAAGGACAGAAAAAGTAGCGAGAA-GAAAAGAAAAGACAGA	1027				
Db	212	GAGGAGAAAGAGAGAGAGAGAAAGAGAGAAAGAGAGAAAGAGAGAAAGAAAGAA	271				
QY	1028	GCCAGCTTCGAGAGAGACAGAGACCCGACAAAGACACACAGGCCCAAGTTGTACGACA	1087				
Db	272	GGATCAGAGAAAGAGGAGGAGAGAGAAAGAGAGAGAGAAAGAAAGAGAGAGAGAGA	331				
QY	1088	CTACGAGAAAGGTGAGCTGCTCTTTGAAAGACCAAGGTTTGATGACTGAGGCATCTTCAGA	1147				
Db	332	AGAGAGAAAGAAAGAGAAAGAGAGAGAGAAAGAAAGAAAGAGAGAGAGAAAGAAAGA	391				
QY	1148	GGAGAAAGTGTGCTCTTTGGCAACGAAAGTGTGATGAGAAAGATGAAAGCCCAACAGA	1207				
Db	392	GGAGAGAGAAAGAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGCAAAAGA	451				
QY	1208	AGTTGTTGCAAGGTCCTCACGTGAGACCGTGGAGAAACAGAGGAGAGACAGAGAGAG	1267				
Db	452	GGAAAGAGGAGGAGAGAGAGAAAGAAAGAAAGAGAGAAAGAGAAAGAGAGAAAGAGA	511				
QY	1268	AGGAGAGGCTGAAGGGGGCGTGG	1290				
Db	512	AGAGAGAGAGAGAGCTGGACTTGG	534				

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2002, 14:13:30 ; Search time 7323 Seconds
(without alignments)

13623.405 Million cell updates/sec

Title: US-09-902-432-3

Perfect score: 6160

Sequence: 1 CGGGCGGGGGGGtagaag.....aaataaacatttgactac 6160

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST.*
1: em_estba.*
2: em_esthum.*
3: em_esthu.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_man.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	751.4	12.2	1153	14 D88456	D88456 D88456 Mous
C 2	682.8	11.1	710	14 BQ200192	BQ200192 UI-R-DQ1-
C 3	634.8	10.3	683	13 BI274459	BI274459 UI-R-CMO-
4	571	9.3	776	13 BG974884	BG974884 602842771
5	552.8	9.0	1269	11 BC015830	BC015830 Homo sapi
C 6	550.4	8.9	746	13 BM230239	BM230239 K0292D03-

C 7	550	8.9	563	9 AI233818	AI233818 EST230506
8	547.6	8.9	661	13 BI652823	BI652823 603298753
C 9	532.2	8.6	710	13 BM119923	BM119923 L0932E09-
10	529.2	8.6	668	10 BB546856	BB546856 BB546856
C 11	512	8.3	703	14 BQ176546	BQ176546 UI-M-DJ2-
12	501.8	8.1	610	9 AI225639	AI225639 u113909.Y
C 13	492.2	8.0	665	13 BM228830	BM228830 K0266G02-
C 14	467.8	7.6	476	9 AI103489	AI103489 EST212778
15	460.2	7.5	573	12 BG670384	BG670384 DRNBAE11
16	459.6	7.5	679	14 BQ443215	BQ443215 UI-M-EV0-
17	458.6	7.4	567	12 BF286252	BF286252 EST450843
C 18	456.6	7.4	623	13 BM226613	BM226613 K0230F11-
C 19	453.8	7.4	825	13 BM226974	BM226974 K0236812-
20	450.6	7.3	530	9 AI156819	AI156819 u144c01.Y
C 21	442.8	7.2	601	13 BM226680	BM226680 K0231G01-
C 22	442.8	7.2	614	13 BM245338	BM245338 K0723D03-
23	439.6	7.1	754	13 BG966814	BG966814 602834309
24	437.8	7.1	745	12 BG663107	BG663107 DRA08D03
25	436.6	7.1	534	10 AW919250	AW919250 EST350554
C 26	434.8	7.1	593	13 BM225990	BM225990 K0220D09-
C 27	430.8	7.0	586	13 BM293486	BM293486 C0305H04-
C 28	429.4	7.0	453	10 AW535062	AW535062 UI-R-BS0-
29	421.4	6.8	641	10 BB665546	BB665546 BB665546
C 30	418.2	6.8	577	13 BM230475	BM230475 K0295H08-
C 31	416.6	6.8	428	9 AI176113	AI176113 EST213692
C 32	415.8	6.8	574	13 BM230613	BM230613 K0298D07-
C 33	414.8	6.7	574	13 BM229218	BM229218 K0275C05-
C 34	412	6.7	589	13 BM244908	BM244908 K0717D04-
35	410.2	6.7	886	12 BF163549	BF163549 601769945
36	408.6	6.6	644	13 BI446672	BI446672 dac81f11
37	406.6	6.6	674	9 AL045541	AL045541 DRP2p434K
38	405.8	6.6	704	13 BI645683	BI645683 603275189
C 39	400.4	6.5	562	14 BQ554344	BQ554344 H4027E12-
C 40	400.2	6.5	483	9 AA963479	AA963479 UI-R-E1-g
C 41	392.6	6.4	552	13 BM124130	BM124130 L0535F11-
C 42	385.2	6.3	599	9 AI317366	AI317366 u113909.X
C 43	384.2	6.2	550	13 BM244784	BM244784 K0715G12-
44	382	6.2	599	14 BQ566352	BQ566352 g158B03.Y
C 45	371.6	6.0	522	13 BM197966	BM197966 C0352A03-

ALIGNMENTS

RESULT 1

D88456

LOCUS

DEFINITION

D88456 Mouse ICR testis (J.Tsushima) Mus musculus

1153 bp mRNA linear EST 01-APR-1998

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Medline

Comment

1 (bases 1 to 1153)

Takeo,M.M., Araki,Y., Matsunaga,A., Yokoi,A., Tsuchida,J., Nishina

,Y., Noraki,M., Tanaka,H., Koga,M., Uchida,K., Matsumiya, Okuyama

,A., Rochelle,J.M., Nishimune,Y., Matsui,M. and Seldin,M.

Mapping of eight testis-specific genes to mouse chromosomes

Genomics 46 (1), 138-142 (1997)

98066773

Contact: Junji Tsuchida

Department of Science for Laboratory Animal Experimentation

Research Institute for Microbial Diseases, Osaka University

3-1 Yamadaoka, Suita, Osaka 565, Japan

Tel: 06-879-8338

Fax: 06-879-8339

Email: tsuchida@biken.osaka-u.ac.jp.

Location/Qualifiers

1. .1153

/organism="Mus musculus"

Features

source

```

/strain="ICR"
/db_xref="taxon:10090"
/clone="A-12"
/sex="male"
/tissue type="testis"
BASE COUNT      367 a      285 c      299 g      201 t      1 others
ORIGIN
Query Match      12.2%; Score 751.4; DB 14; Length 1153;
Best Local Similarity 81.1%; Pred. No. 1.6e-159;
Matches 975; Conservative 0; Mismatches 157; Indels 70; Gaps 6;

QY 4081 GAGGAGGCGCTTCTCCCAAGCAAGAGGTTGCTTAAGGTTCAAGTTCAA 4140
Db 1 GGGGAGGCGCTTCTCCCAAGCAAGAGGTTGCTTAAGGTTCAAGTTCAA 60
QY 4141 GCGTGAACAACAGTCACTCAACAGAGAGCTGTGGAAGAGTCAAGAGGTTG 4200
Db 61 GCTGAGACACACAGTCACTCAACAGAGAGCTGTGGAAGAGTCAAGAGGTTG 120
QY 4201 TGATTTCAGAGACAGTGAAGAGTGTGAGTGTGACACATTATTAACAGCTGAGA 4260
Db 121 CAATTTCAGAGATGATGAAGAGTGTGAGTGTGACAGTGTGCAATCATTAACAGCTGAGA 180
QY 4261 AGTCTCTGCAACGGGTGGCCCTGGACTTTCAGCATGCAAGAGACACGTTG 4320
Db 181 AGCTCTCGAAGACCGGTGGCTTACGGGACTTTCAGCATGGAAGAGACCGTGGCCGAG 240
QY 4321 GGCGTGAAGTCAAGAGATCCATCCCATATGTAAGTCACTCTCTCTGTAAGACACC 4380
Db 241 NGCTGAAGTCTGAGCAGAGTCCATCCCATATGTAAGTCACTCTCTCTGTAAGATCC 300
QY 4381 TACATCTGACCTTACAGAGAGAAATTAAGGCAATCCAGAGAGCGATCAAGAGAG 4440
Db 301 TACATCTGACCTTCAAGAGAGAGAGAGGCAATCCAGAGAGAGATGAGTGAAGATA 360
QY 4441 ACAAGCCAGATGCTGTCTGTATGCTGAGCGAGAGAGAGTACAGCATCCGAAGAGTCC 4500
Db 361 ACAAGCCAGATGCTGTCTGTATGCTGAGCGAGAGAGTACAGCATCCGAAGAGTCC 420
QY 4501 TCAAGGCTGAACCTGAGATCCTGGAATCTGAGATTAAGAGCAAGATTTGTCTGAACG 4560
Db 421 TCAGGCTGAACCTGAGATCCTGGAATCTGAGATTAAGAGCAAGATTTGTCTGAAGTG 480
QY 4561 TCATTTCAGACAGCGGTGACAGATTCGACAGTCAAGAAACAGCCCCGAACTCATGCTT 4620
Db 481 TCATTCAGACAGCGGTGACAGATTCGACAGTCAAGAAACAGCCCCGAACTCATGCTT 540
QY 4621 ATGATTTCACAGACCAAGGTTCTGCAATGCAAGGCTTGACAGCGAGGACCCCAAGATGCT 4680
Db 541 CTGATTTTCAGATCAAGATGCTTCT-----GTGATGTCAGCCAGGAGACCAACATATGCT 593
QY 4681 GGACAAAATGAAG-----ATGCCAAGATGAACACCCAGTGGCCGAGC 4725
Db 594 GGACAAAATGAAGCGACCTTCAAGTCTCCCCCAAGATGGAACATCTAGTGGCTAGC 653
QY 4726 CCAGAGAGACTTTCAGAGTCTGACCGTTCTGAGAGCATGGGCTCAGCCTCGAAATGCT 4785
Db 654 CCAGAGAGACTTTCAGAGTCTGATGATGTTCTGAAGCATGACCAAGGCTTCAGAAATGAT 713
QY 4786 TGCCGCGCTTTCAGATGGAAGCGCGGTGTCAAAGTAAAGCATTTAGAGAGTCTCTCTCA 4845
Db 714 CACCAAGCTTTCAGATGGAAGGTGCAAGTGTCAAAGAAATGTTGGAAGTCTCTCTCTCA 773
QY 4846 ACCCAAGATTCAAAAGAGCATGCTGCTGATGAGGCGCTCAGTCCAAAGCTTAGGCCAGG 4905
Db 774 GTGCAAAAGT-GAAAAGGAACATGCTGCTGACGGCCCCCAGAGCAAAAGCTTAGGCCAGG 832
QY 4906 CAGAGGCGAGTCTCTGGAACCTTAACCAAGAAATCCCAAGACCAACGAGCAACCAAGC 4965
Db 833 CAGAGGCGAGTCTCTGGAATCTTAACCAAGAAATCCCAAGACCAACGAGCAACCAAGC 892

```

```

QY 4966 TAACGAGAGAGGCGATCCCAAGAGTTGAGTTCAGAGAGAAAGATGTTCAACAAGT 5025
Db 893 TAACGAGAGAGGAGATGCTCTGAAGAG----- 922
QY 5026 CAGTCAAGAGACAGAGGCGCCAGGACAGAGAG--GACCTGACAGAGCCAAAGGAGAGCC 5082
Db 923 -----AATGAACAGAGGCCAGACAGAGAGAGAGAGCTTACAGAGCCAAAGGAGAGCC 976
QY 5083 TGGCAGAAATCTTAAGATGTTAGTGTCTCATTTGTCATCTGTAAAGACCAAGATGTAAGAAC 5142
Db 977 TGACAGAAATCTTAAGAGTGTGTTGTCAT-----TGTAAACATGAATGTAAGATG 1028
QY 5143 AAGTCAAGAACAGATGCTGCTGTGGACCTTGAGCCAAAGATTCAGAGCCCATGAG 5202
Db 1029 AAGTCAAGAACAGATGCTGCTGTGGACCTTGAGCCAAAGATTCAGAGCCCATGAG 1068
QY 5203 ATCCAGAGAGAGGCGCCGCTCCAAATGTTCCAGCCAGTAGAGCAACCCGCAATCTGAG 5262
Db 1089 GTGACAGAGAGAGGCGCCGCTCCAAATGTTCAACCTGACAGAGCAACCCGCAATCTGAG 1148
QY 5263 GC 5264
Db 1149 GC 1150

```

RESULT 2

B0200192/c 710 bp mRNA linear EST 02-MAY-2002

LOCUS

UI-R-DQ1-clu-e-23-0-UI.s1 UI-R-DQ1 Rattus norvegicus cDNA clone

DEFINITION

UI-R-DQ1-clu-e-23-0-UI 3', mRNA sequence.

ACCESSION

B0200192

VERSION

B0200192.1 GI:20416657

KEYWORDS

EST.

SOURCE

Norway rat.

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 710)
Bonaldi,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

JOURNAL

COMMENT

MEMLINE

CONTACT: Soares, MB

Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
normalized rat cell line R3327-5a library cDNA library Preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com)
Seq primer: M13 Forward

POLYA=yes.

FEATURES

source

Location/Qualifiers

1..710

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-DQ1-clu-e-23-0-UI"

/clone_lib="UI-R-DQ1"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7AD-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-DQ1
library is a normalized Rat Cell Line R3327-5A library


```

|||||
Db 504 AAGTCTTTTAAAGTTTACTATGATCTGTGGGCTTCTAGTCTCTGAAAGTGG 445
Qy 5737 TTGTTTCTATGACAGGAGCTCGAATAAATAAAAACCCATTTTGAACATCCAGATG 5796
Db 444 -TGTTCCTATGACAGGAGCTCGAATAAATAAAAACCCATTTTGAACATCCAGATG 386
Qy 5797 TCCCAATATATACAGATTTTTCCTTTTCCCTTTTGTCTAATCCAGTCCAGTTGAAAGAA 5856
Db 385 TCCCAATATATACAGATTTTTCCTTTTCCCTTTTGTCTAATCCAGTCCAGTTGAAAGAA 326
Qy 5857 GTCCTCTGTCGATGATTAAGCCCTGTCTCTTAATGATGACAAATGAGTGTGCTTA 5916
Db 325 GTCCTCTGTCGATGATTAAGCCCTGTCTCTTAATGATGACAAATGAGTGTGCTTA 266
Qy 5917 AGCCATGAGATGTTTCTTAATGACAGAAAGAACTGTGTAGCGTTTGTGATGATGCTC 5976
Db 265 AGCCATGAGATGTTTCTTAATGACAGAAAGAACTGTGTAGCGTTTGTGATGATGCTC 206
Qy 5977 TTCTATGCTGACCGAATTCATATGCAATGCAAGTGAATCCTGTCTTTTACAGATGTA 6036
Db 205 TTCTATGCTGACCGAATTCATATGCAATGCAAGTGAATCCTGTCTTTTACAGATGTA 146
Qy 6037 TTTTATGATGACTGGAGTTGTCTGTATATATCTGTGCCCCCTTTTAAAGAAATGT 6096
Db 145 TTTTATGATGACTGGAGTTGTCTGTATATATCTGTGCCCCCTTTTAAAGAAATGT 86
Qy 6097 TGCATATATGCTCTTGTGATTAATGTGATTTGACAACTGATTTTAAATAACATTTTGA 6156
Db 85 TGCATATATGCTCTTGTGATTAATGTGATTTGACAACTGATTTTAAATAACATTTTGC 26
Qy 6157 CT 6158
Db 25 CT 24

RESULT 4
Bg974884 776 bp mRNA linear EST 12-JUN-2001
LOCUS 602842771F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4978744 5',
DEFINITION mRNA sequence.
ACCESSION Bg974884
VERSION Bg974884.1 GI:14362521
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 776)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10975 row: m column: 17
High quality sequence stop: 753.
Location/Qualifiers
1. 776
/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone_1b="IMAGE:4978744"
/clone_1lb="NCI CGAP Mam4"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"

```

```

/lab host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI. Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."
```

BASE COUNT 278 a 181 c 216 g 101 t

ORIGIN

Query Match 9.3%; Score 571; DB 13; Length 776;
Best Local Similarity 86.3%; Pred. No. 1.3e-118;
Matches 677; Conservative 0; Mismatches 90; Indels 12; Gaps 4;

```

Qy 327 GAAAAAGCCGAGTTGAAGAAATGGCGCCCACTCCACAGCTGTTGAAGATATCAAAAG 386
Db 1 GAAAAAGCCGAGTTGAAGAAATGGCGCCCACTCCACAGCTGTTGAAGATATCAAAAG 60
Qy 387 GATGGCAGAGAGAGACATCGAATAATTAAGACAGATCCCTGCTTCAGAAAACATGTG 446
Db 61 GACGAGCAGAGAGAAACCCGAAATTAATGACAGATCCCTGCTTCAGAGACAAATGTG 120
Qy 447 GAAGAAATGTATACAGCTCTGAGTCCAGGCTATATGATGTGTGCTTCAGAAAATATT 506
Db 121 GAAGAAATGTGCGAGGCTGTGAGTCCCAAGCTATATGACGTGCGCTTCAGAAAGTATT 180
Qy 507 AATTTGTGTTTAAATTTACCGTGAAGAAAGATTAATAATGAAAATCGATATCTGT 566
Db 181 AATTTGTGTTTAAATTTACCGTGAAGAAAGATTAATAATGAAAATCGATATCTGT 240
Qy 567 CAATCTACTCTGTCAAGAAAGATGAAGGCGAAGGCGAAGAGCTGTGCGAGCTGCA 626
Db 241 CAGTCTACTCTGTCAAGAAAGATGAAGGCGAAGGCGAAGAGCTGTGCGAGCAGGA 300
Qy 627 GACCAACAGAGAGCCAGTGTGAGAGCTGCCGTGAGAGTCAAGATCCAAAGAAATGTAG 686
Db 301 GACCAACAGAGAGCCAGTGTGAGAG--CGTGGCGAATCAGCATCCAAAGAAATGTAG 357
Qy 687 CTGAAGCAATCCACAGAGAGAGAGAGGCACTCTGAAGCAAGAAACAGACAGACAGAA 746
Db 358 CTGAAGCAATCCACAGAGAGAGAGAGGCACTCTGAAGCAAGAGCAAGACAGACAGAA 417
Qy 747 ATCCCCCTTCAAGCCGATCTGATCAAGCGGCTGAGAAAGGCCAAAGATGAAGAGAA 806
Db 418 ATCCCCCTTCAAGCCGATCTGATCTGATCAAGCGGCTGAGAAAGGCCAAAGATGAAGAA 477
Qy 807 GAAAAACAAGAGAAAGAGCCCAAGTCCCGAGATCCCGAGAGCCCAAGTCAACAGT 866
Db 478 GAAAAACAAGAGAAAGAGCCCAAGTCCCGAGATCCCGAGAGCCCAAGTCAACAGT 537
Qy 867 GAGACAAATCTTCTTCAAGAAATTTCTTCACTCACTGCGTGGCTGGCTGGCGCAAGAG 926
Db 538 GAGACAAATCTTCTTCAAGAAATTTCTTCACTCACTGCGTGGCTGGCTGGCGCAAGAG 597
Qy 927 ACCAGCTTCAAGAAATCAAAAAGAGAGATGATCGAAATCGCCGAGAGAAAGAGAGCA 986
Db 598 ACCAGCTTCAAGAAATCAAAAAGAGAGATGATGGAATCTTCCGAGAGAGACAGAGAGCA 657
Qy 987 GAGGCAAGAAATGAGACGAGAGAGAAAGG---AAAAAGACAGAGCCAGCTCGAGAGAG 1043
Db 658 GAGGCTGAAAAAAGTGAACGAGAGAGCAAGGGAGCAAGAGACAGAGCCAGCCGAG 717
Qy 1044 CAGGAG-----CCGGCAGAGACAGAGACAGAGCCAGGTTG-TAGCAGATCTACAGAA 1096
Db 718 GAGCAGAGAGAGCTGCAAGAGAGGCAAGAGCCAGGCTCAGAGATGACCGCAGATGAGAA 776

RESULT 5
BC015830 1269 bp mRNA linear HTC 17-OCT-2001
LOCUS BC015830
DEFINITION Homo sapiens, clone IMAGE:4691862, mRNA.
ACCESSION BC015830
VERSION BC015830.1 GI:16198369

```

KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS 1 (bases 1 to 1269)
TITLE Strausberg, R.
JOURNAL Direct Submission
REMARK Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 38 Row: k Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4826749
This clone has the following problem: frame shifted.
FEATURES
source Location/Qualifiers
1. 1269
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4691862"
/tissue type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
BASE COUNT 419 a 311 c 365 g 174 t
ORIGIN
Query Match 9.08; Score 552.8; DB 11; Length 1269;
Best Local Similarity 69.8; Pred. No. 1.8e-114;
Matches 894; Conservative 0; Mismatches 287; Indels 99; Gaps 7;
Qy 10 GGAGTAGAAGAGCACTAGCCATGGCGCAGGAGTTCACCGAGCAGCGGAGCCCG 69
Db 2 GGGGGAGAGTGGGAGGAGCCATGGCGCGGAGCTCCACCGAGCAGCGCGCGG 61
Qy 70 AGCAGCCG---GCGGGAGCAGACCGCGAGCGTGTGCTCAGTGGCCATGGGCGCG 126
Db 62 AGCAGCCGCGGAGGAGTCCACCGCGGTGAGCCGCGAGCCCGCGCGGCGCCCT 121
Qy 127 CAGCTGAAGCTCCGGA-----GCAGCTGGAGACCCCGCGAGCCCGCGCGCA 177
Db 122 CGGCCGAGCGCGCGCGAGACCAACCGCGAGCCCGCGCATCGCTCGGACCCCGCA 181
Qy 178 CCAAGCTCCACAGAGAAGTGGCAGCTGTCTGTCAACGGCGGTAGCTGAACAAGAG 237
Db 182 CCAAGCTCCTACAGAGAGTGGTCAAGTGTCCACCATCAATGGCGGTAGCTGAGCAAGATG 241
Qy 238 ATGTCCATGTCCAAAGAGGAAACACGAGGG----- 268
Db 242 AGCTCAGCTCCAGGAGGTGACCTTAATGGCCAGAAAGAGCCCTGACGGTCAAGGAG 301
Qy 269 -----GCAGGAGGAAGTGTGTGATGAGGATGTTGGACGAGCAGAGTCAG 315
Db 302 CCCTAAACAGCCAGGAGGAGGAAGATGATGTACAGAGGTTGGACAGAGACTCTG 361
Qy 316 AAGATGTGAGAGAAAGACCGAGTTGNAGAAATGGCGGCCCACTCCACAGCTTTGAG 375
Db 362 AAGATGTGAGCAAGAGACTCCGATAAAGAGATGGCTACTAAGTCAGCGGTTGTTACG 421

Qy 376 ATATCAAAAGGATGGCAGGAGGAGACATCAGAAATAATTGAACAGATCCCTGCTTCAG 435
Db 422 ACATCAGAGATGATGGCAGGAGGAGACACCCGAAATAATCGAACAGATTCCTTCTTCAG 481
Qy 436 AAACAAATGTGGAAGAATGGTACAGCTCTGAGTCCAGGCTTAATGATGTTGGCTTCA 495
Db 482 AAAGCAATTTAGAGAGACTAACCAACCCACTGAGTCCCAGGCTTAATGATTTGATTTA 541
Qy 496 AGAAAGTATTAAATTTGTTGGTTTTAAATTCACGGTGAAGAAGGATAAAAAAAGT 555
Db 542 AGAAGTGTTTAAAGTTTGGCTTTAAATTTCACTGTGAAAAAGGATAAGACAGAGAAGC 601
Qy 556 CAGATCTGTCCAACTACTCTCTGTCAGAAAGGATGAAGCGGAAGGGGCGAGACCTCTG 615
Db 602 CTGACACTGTCCAGCTACTCTGTCAGAAAGATGAAGGGGAGGAGCAGCA----- 654
Qy 616 TCGAGCTGTGAGACCCACAGGAGCCAGTGTGGAGACTCGCGTCGGAGAGTCAGCATCCA 675
Db 655 -GGGGCTGGGACCAACAGGACCCCGAGCTTTGGGCT-----GGAGAGCAGCATCCA 706
Qy 676 AAGAAAGTGTGAGCTGAAGCAATTCACAGAGAGCAAGGAGCCCTGAAGCAAGACAGA 735
Db 707 AAGAAACCGAACCCCAACAATCTACAGAGAAACCCGAGAGACCTCGAGCGTGAGCAA 766
Qy 736 GCAGCAGAAATCCCTCTCAAGCCGAAATCTGATCAAGCGGTGAGGAAGACCCAAAG 795
Db 767 GCCACGAGAAATTTCTCCCGCCAGCCGAAATCTGGCCAA---GCAGTGGAGGAATGCAAG 823
Qy 796 ATGAAGGAGAGAAACAGAGAAAGAGCCACCAAGTCCCGAGATCCCGAGCAGCC 855
Db 824 AGGAGGAGAGAGAAACAAAGAAAGAACTAGCAAGTCTGCAGAAATCTCCGACTAGTC 883
Qy 856 CAGTCAACAGTGTGAGCAACAATCTTCTTCAAGAAAGTTCTTCACTCAGGTTGGGCCCGCT 915
Db 884 CCGTACAGTGTGAACAGGATCAACCTTCAAAAAATTTCTTCACTCAGGTTGGGCCCGCT 943
Qy 916 GCGCAAGAGACACAGCTTCAAGAAATCAAAAGAGGATGATCTGGAACCTGCCGAGAGA 975
Db 944 GCGCAAAAAAGACCAAGTTTCAGGAAGCCGAGGAGGATGAAGTGAAGCTTCAGAGAAGA 1003
Qy 976 GAAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035
Db 1004 AAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1063
Qy 1036 CCGAGGA-----GCAGGAGCCGCGAGAGACACAGAC 1068
Db 1064 CCGAGAACTGACCGCTCCGAGCAAGCCACCCACAGGAGCGGCGAGAAAGTCCCAAG 1123
Qy 1069 AGCCAGGTTGTGAGCAGACTACGAGAAAGTGGAGTGCCTTTGGAAAGACAGGTTGGTG 1128
Db 1124 AGCCCGGTTATCAGCTGAATATGAGAAAGTGAAGTGCCTTCAGAGGAGCAAGTCAGTG 1183
Qy 1129 ACCTGGAGGATCGTCAGAGGAGAGTGTCTCTTGGCAACGGAAGTGTTCATGAGA 1188
Db 1184 GCTCGCAGGAGACCTTCTGAAGAGAAACCTGCTCGTGGCGACAGAAAGTGTTCATGAAA 1243
Qy 1189 AGATGGAAGCCCAAGAA 1208
Db 1244 AAAAAAAAAAAAAAAAAA 1263

RESULT 6
BM230239/C
LOCUS BM230239 746 bp mRNA linear EST 31-JAN-2002
DEFINITION K0292D03-3 NIA Mouse Unfertilized Egg cDNA Library (Long) Mus
musculus cDNA clone K0292D03 3', mRNA sequence.
ACCESSION BM230239
VERSION BM230239.1 GI:17793419
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 746)	Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus			
Piao,T., Kargul,G.U., Dudekula,D.B., Qian,Y., Luo,A., Stegg,C.A., Martin,P., Abba,K., Tanaka,T. and Ko,M.S.H.				
Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library (long)				
Unpublished (2001)				
Contact: Dawood B. Dudekula				
Laboratory of Genetics				
National Institute on Aging/National Institutes of Health				
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA				
Email: cdna@igsun.grc.nia.nih.gov				
Plate: K0292 row: D column: 03				
Seg primer: -21M13 Forward				
High quality sequence stop: 746				
POLYA=Yes.				
FEATURES				
source				
1. . 746				
/organism="Mus musculus"				
/strain="C57BL/6J"				
/db_xref="niaEST:K0292D03-3"				
/db_xref="taxon:10090"				
/clone="K0292D03"				
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (long)"				
/tissue_type="Unfertilized Egg"				
/lab_host="DH10B"				
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2				
NotI; Mouse cDNA project by the Laboratory of Genetics,				
National Institute on Aging (NIA), Intramural Research				
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is				
a long-cytoplasmic enriched cDNA library (Ref. Genome Res.				
11: 1553-1558 (2001). [PMID: 11544191]. Total RNAs were				
extracted from a pool of 1488 unfertilized eggs				
Double-stranded cDNAs were synthesized with an Oligo(dT)				
primer (Invitrogen):				
5'-TGACAGTCTTCAGATCGGAGCGCGCCCTTTTCTTTTCTTTT-3'],				
treated with T4 DNA polymerase, and purified by				
ethanol-precipitation. The cDNAs were ligated to				
lone-linker lL-SalI, purified by phenol/chloroform, and				
separated from free linkers by Centricon 100. Then, the				
cDNAs were amplified by long-range high fidelity PCR using				
Ex Tag polymerase (Takara) with a primer Sal4-S. The				
products were purified by phenol/chloroform and Centricon				
100. The cDNAs were digested with SalI and NotI enzymes				
and cloned into SalI/NotI site of pSPORT1 plasmid vector.				
The DH10B E. coli host was transformed with the ligation				
mixture by the standard chemical method. The average				
insert size is about 2.5 kb. The library was constructed				
by Yulan Piao (NIA)."				
BASE COUNT				
261 a 137 c 147 g 201 t				
ORIGIN				
Query Match				
Best Local Similarity 90.4%; Score 550.4; DB 13; Length 746;				
Matches 678; Conservative 0; Mismatches 61; Indels 11; Gaps 8;				
5412 TAAAGTCCGCGCTTCACACCTGGAAACCAAT-TCGGCACTACCTAGTCCACTTCGAAA				
5470				
Db 746 TAAAGTCCGCGCTTCACACCTGGAAACCAATGCGCAATACCTATCCACTTCGAAA				
687				
5471 CTGAGACATCCCTCTTATGATTTATATGATATGTTTATGATAGTCTCTCTGTAACCT				
5530				
Db 686 CTGAGACCTCCCTCTTATGATTTATATGATATGTTTATGATAGTCTCTCTGTAACCT				
627				
5531 ATTGTATA-TTTTTTCTAAGCTTTAAGACATGCTTTTGTATTATGCAATATATAAG				
5589				
Db 626 ATTGTATAATTTTTTTCTAAGGTTTAAAGACATGCGCTTTGTATATAGCATATATAAG				
567				
5590 GGTGTGAGCCATAGGAGAGCGCTTGGAAAAGCTCCAGGCTCAACCTGAACCTGACACAA				
5649				
Db 566 GGTGTGACCACTAGGAGAGCGCTTGGAAAAGCCCCAGGCTCAACCTGAACCTGACACAA				
507				

Accession	Source	Organism	Version	KeyWords	Accession	Source	Organism	Version	KeyWords
5650					5705				
5660					5705				
506					448				
5706					5765				
447					388				
5766					5825				
387					329				
5826					5885				
328					269				
5886					5945				
268					210				
5946					6005				
209					151				
6006					6065				
150					91				
6066					6124				
90					31				
6125					6154				
30					1				

RESULT 7
 A1233818/c
 LOCUS
 DEFINITION
 A1233818 563 bp mRNA linear EST 31-JAN-1999
 E57230506 Normalized rat lung, Bento Soares Rattus sp. cDNA clone
 RUCR55.3' end, mRNA sequence.
 ACCESSION
 A1233818
 A1233818.1 GI:3817698
 VERSION
 A1233818
 KEYWORDS
 EST.
 SOURCE
 Rattus sp.
 ORGANISM
 Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 563)
 Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
 Kerlavage,A.R. and Adams,M.D.
 Rat Genome Project: Generation of a Rat EST (R5E) Catalog & Rat
 Gene Index
 Unpublished (1998)
 Other ESTs: TC58449
 Contact: Lee, NH
 The Institute for Genomic Research
 9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-3529
 Fax: (301)-838-0208
 Email: nhlee@igir.org
 Seq primer: M13-21.
 Location/Qualifiers
 1..563
 /organism="Rattus sp."
 /db_xref="ATCC (inhost):2039638"
 /db_xref="taxon:10118"
 /clone="RUCR55"
 /clone_lib="Normalized rat lung, Bento Soares"
 /note="Organ: lung; Vector: pT713Pac; Site_1: EcoRI;
 Site_2: NotI"

BASE COUNT	191 a	107 c	110 g	155 t	
ORIGIN					
Query Match	8.9%; Score 550; DB 9; Length 563;				
Best Local Similarity	99.1%; Pred. No. 7.2e-114;				
Matches 553; Conservative	0; Mismatches 5; Indels 0; Gaps 0;				
QY	5601	ATAGCGAGCTTTGAAAGCTCGAAGCTCAAGCTCAACTGTGTAACTGTGAGCAAGCAAAAGATAACATT	5660		
Db	563	AAAGCGAGCTTTGAAAGCTCGAAGCTCAAGCTCAACTGTGTAACTGTGAGCAAGCAAAAGATAACATT	504		
QY	5661	CTGCGCAAGAGACAGACAGCTTTTAAAGTTTAAAGTTTAACTGTGTAGCTTAGATCTGTGGGCTTCT	5720		
Db	503	CTGCGCAAGAGAGACAGACAGCTTTTAAAGTTTAAAGTTTAACTGTGTAGCTTAGATCTGTGGGCTTCT	444		
QY	5721	AGTCCTCTGAAAGTGTGTGTTTCTTATGACAGCGAGCTCAGAAATAAAGCAACCCCAATTT	5780		
Db	443	AGTCCTCTGAAAGTGTGTGTTTCTTATGACAGCGAGCTCAGAAATAAAGCAACCCCAATTT	384		
QY	5781	TGAACATCCAGGATGTCCCAATATTACCATGATTTTTCCTCCCTTTTGTCTAATCCAG	5840		
Db	383	TGAACATCCAGGATGTCCCAATATTACCATGATTTTTCCTCCCTTTTGTCTAATCCAG	324		
QY	5841	TCCAGGTTGGAAGAAAGTCTCTCTGTGTCAAGATTAAAGCCCTGTCTCTTAATGATATGGA	5900		
Db	323	TCCAGGTTGGAAGAAAGTCTCTCTGTGTCAAGATTAAAGCCCTGTCTCTTAATGATATGGA	264		
QY	5901	CAATAGTGTGCTTAAGGCCATGAGATGTTTCTTAATGCAAGAGGAATCTGTTGTACGT	5960		
Db	263	CAATAGTGTGCTTAAGGCCATGAGATGTTTCTTAATGCAAGAGGAATCTGTTGTACGT	204		
QY	5961	TTTTTGAATGTACTCTTCTATGCTGACCGAATTCATATGACAGATCGAAGTGTGCTG	6020		
Db	203	TTTTTGAATGTACTCTTCTATGCTGACCGAATTCATATGACAGATCGAAGTGTGCTG	144		
QY	6021	TTCTTTACAGATGATTTTATGATAGACTGGAGTTTGTCTGTGTTATATCTGTCCTCT	6080		
Db	143	TTCTTTACAGATGATTTTATGATAGACTGGAGTTTGTCTGTGTTATATCTGTCCTCT	84		
QY	6081	TCCTTAAGAACAAATGTTGCAATATGTTCTTTGGATAAATGTTGATTTGCAAACTGATTT	6140		
Db	83	TCCTTAAGAACAAATGTTGCAATATGTTCTTTGGATAAATGTTGATTTGCAAACTGATTT	24		
QY	6141	AAATAAACATATTGACT	6158		
Db	23	AAATAAACATATTGCTCT	6		
RESULT 8					
BI652823					
LOCUS	603298753F1 NCI_CGAP_Mam3 Mus musculus cdna clone IMAGE:5339244 5',				
DEFINITION	mRNA sequence.				
ACCESSION	BI652823				
VERSION	BI652823.1				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (bases 1 to 661)				
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapbs-@email.nih.gov				
	Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	http://image.llnl.gov				
Plate: LLAM1861	row: j column: 13				
High quality sequence stop: 661.					
Location/Qualifiers					
1. .661					
/organism="Mus musculus"					
/strain="129,C57BL/6J,FVB/N"					
/db_xref="taxon:10090"					
/clone="IMAGE:5339244"					
/clone_lib="NCI CGAP Mam3"					
/tissue_type="tumor, gross tissue"					
/dev_stage="10 months"					
/lab_host="DH10B"					
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;					
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.					
Library constructed by Life Technologies. Investigators					
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH					
Reference for transgenic model: Xu et al., Nature Genetics					
22, 37-43 (1999)." 98 t					
BASE COUNT	183 a	179 c	201 g		
ORIGIN					
Query Match	8.9%; Score 547.6; DB 13; Length 661;				
Best Local Similarity	90.2%; Pred. No. 2.6e-113;				
Matches 597; Conservative	0; Mismatches 64; Indels 1; Gaps 1;				
QY	2656	CAGTCATTTGATGGACAGGGCAGTCAACAGTGTGCGAAGAGCGGTCTCTTCGTGGATAT	2715		
Db	1	CGGTCTATTGATGGACCA-GGCAGTCACCAAGTCCGAAGAGCGGTCTCTTCGTGGATAT	59		
QY	2716	CCGCTTCCGTAACAGAACCTTTGAACACACAGGGGAGAGCCATGCCACCTGTTGAAG	2775		
Db	60	CTGCTTCCATGACAGAACCTTTGAGCACGACAGAGGAGTGGCCACACCCCTGTTGGAG	119		
QY	2776	AGGTCACTGAAAAAGACATCAATTCAGAGAAACTCTCTGTGCTCACCCAGAGTTACAG	2835		
Db	120	AGGTCACTGAAAAAGACATCACTGCGAAGCAACTCTCTGCACCTGCCACAGACTTTACAG	179		
QY	2836	AGGTAAAGATGCCCATGACGACATGGTCAACAGTGAAGTGGATTTTCCCTCAGAAAGCTG	2895		
Db	180	GGGGCAAGATGCCCATGACGACATAGTCAACAGTGAAGTGGATTTTACCTCAGAAAGCTG	239		
QY	2896	TGACAGCCACAGAGACCTCAGAGGCTCTCGTACTGAGAAAGTTACGAGAGATCGGGGG	2955		
Db	240	TGACAGCCGAGAAACCAACAGAGGCGCTCCGCGCTGAAAGAACTTACCGAAGCATCAGGGG	299		
QY	2956	CCGAAGACACAGACATGCTGTCCGAGTTTCCAGCTGACTGCTCCCGACACCA	3015		
Db	300	CAGAAGACACAGACATGCTGTCTGCAGTTTCCAGCTGTCCGACTCCCGACACCA	359		
QY	3016	CAGAGGAAGCCACCCAGTTTCCAGAGGTAGAGATGGTGTGCTAGATACAGAAAGAGG	3075		
Db	360	CAGAGGAAGCCACCCAGTTTCCAGAGGTAGAGGTGTCATGATACGGAAGACAGG	419		
QY	3076	AGGCGCAGACGAGGCGATCTCTCAAGCCCTTCAGACAAAGGTGAAAGAGAGTCCAGG	3135		
Db	420	AGGCGCAGACGAGGCGCTCTCCAAGCCCTTCAGACAAAGTGAAGAGAGTCCAGG	479		
QY	3136	TGCTGCAACCCAGACATGTCGACAGAACCGGGTCAAAAGCACTGGAGAGGTTGAGGAGG	3195		
Db	480	TGCTGCAACCCAGACATGTCGACAGAGCAGGCGCGAAGCACTGGAGAGGTTGAGGAGG	539		
QY	3196	TAGAGGAGGACTCCGAGTGTCTGGCTTCGAGAAAGAGAGGAGCTTTATCCGAAAGGAC	3255		
Db	540	TAGAGGAGGACTCCGAGTGTCTGGCTTCGAGAAAGAGAGGAGTGTGTGTCGCGAAGGAC	599		
QY	3256	CCGTGCAAGGAGCTGGAGCTGAGCATCTTCCACAGGCTCTGAGACTGGACAGGCTACTC	3315		
Db	600	CCGTGCAAGGAGCTGAAACTGAGCATCTTTCACAGGCTCCGAGAGCTGTACAGGCTACCC	659		
QY	3316	CA	3317		
Db	660	CA	661		

RESULT 9
 BM119923/c 710 bp mRNA linear EST 01-FEB-2002
 LOCUS L0932E09-3 NIA Mouse Newborn Kidney cDNA Library (Long) Mus
 DEFINITION musculus cDNA clone L0932E09 3', mRNA sequence.
 ACCESSION BM119923
 VERSION BM119923.1 GI:17087949
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 710)
 Pao, Y., Kargul, G.D., Dudekula, D.B., Qian, Y., Pantano, S., Lim, M.K.
 and Ko, M.S.H.
 TITLE Systematic Analyses of NIA Mouse Newborn Kidney cDNA Library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 Plate: L0932 row: E column: 09
 Seq primer: -21M13 Forward
 High quality sequence stop: 710
 POLY=A=yes.

FEATURES
 source
 Location/Qualifiers
 1..710
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="tae:EST:L0932E09-3"
 /db_xref="taxon:10090"
 /clone="L0932E09"
 /clone_lib="NIA Mouse Newborn Kidney cDNA library (Long)"
 /cvsuc_type="Newborn Kidney"
 /dev_stage="Newborn"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11:1553-1558 (2001). [PMID:1154199]). In brief , double-stranded cDNAs were synthesized with an oligo(dT) primer (Invitrogen: 5'-GACTAGTCTAGATCGGAGCGGCCGCCCTTTT-3') from 26 micrograms of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker Lf-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes, and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Pao (NIA)."

BASE COUNT 252 a 130 c 137 g 191 t
 ORIGIN

Query Match 8 6%; Score 532.2; DB 13; Length 710;
 Best Local Similarity 90.5%; Pred. No. 8e-110;
 Matches 636; Conservative 0; Mismatches 58; Indels 9; Gaps 6;

Oy 5459 CCACCTTCAACAGGAGATCCTCTTATGATATATATGATATGATAGTCT 5518
 |||||
 Db 710 CCACCTTCAACAGGAGATCCTCTTATGATATATATGATATGATAGTCT 651

Oy 5519 CCTCTGACTATTGTATA-TTTTCTTAAAGTTAAGACATGCTTTTGTATATG 5577

Db 650 CCTCTGACTATTGTATATTTTTTTCTTAAGTTTAAGACATGCTTTGTATATG 591
 |||||
 Oy 5578 CAATATATTAACGGGTGTGACCCATAGCAGCGCTTTGAAAAAGCTCAAGCTCAACTGTA 5637
 |||||
 Db 590 CAGATATATTAATGGGTGTGACCCATAGCAGCGCTTTGAAAAAGCTCAAGCTCAACTGTA 531
 |||||
 Oy 5638 ACCCTGACCAACAGG---ATTAATTCCTGCGCAAGAAAGACAAAGCTTTTAAAGTT 5693
 |||||
 Db 530 ACCCTGACCAACAGGAAATTAATTCCTGCGCAAGAAAGCTTTTAAAGTT 472
 |||||
 Oy 5694 TACGATGCTTAGATCTGTGGGCTTTAGTCCCTGAAAGGTTGTTTCCCTATGACAA 5753
 |||||
 Db 471 TACGATGCTTAGATCTGTGGGCTTTAGTCCCTGAAAGGTTGTTTCCCTATGACAA 412
 |||||
 Oy 5754 GCGAGCTCAGAAATTAACCCCATTTTGAACATCCAGAGTGTCCCATATTAACATGA 5813
 |||||
 Db 411 GTGAGCTCAGAAATTAACCCCATTTTGAACATCCAGAGTGTCCCATATTAACATGA 352
 |||||
 Oy 5814 TTTTTCCTCCCTTTTTCCTTAATCCAGTCCAGTTGAAAGAAATCTCTCTGTGTCAGA 5873
 |||||
 Db 351 CATTTTTCCTTTTTCCTTAATCCAGTCCAGTTGAAAGAAATCTCTCTGTGTCAGA 292
 |||||
 Oy 5874 TTAAGCCCTGCTCTTAATGATATGACAAATGAGTGTGCTTAAGCCCATGAGATGTTTC 5933
 |||||
 Db 291 TTAAGCCCATCTCTTAATGATATGACAAATGAGTGTGCTTAAGCCCATGAGATGTTTC 233
 |||||
 Oy 5934 CTAATGACAAAGAAATCTGTTGACGTTTTTTTGTATGTACTCTTCTAAGTCGACCGAA 5993
 |||||
 Db 232 CTAATGACAAAGAAATTTGTTGT-TACTTTTGTGATTTGTAAGTGTGACCGAA 174
 |||||
 Oy 5994 TTCAATGACAGATGAAAGTGAAGTCTGTTCTTTACAGATGATTTTGTATGATACGTA 6053
 |||||
 Db 173 TTCAATGACAGATGAAAGTGAAGTCTGTTCTTTACAGATGATTTTGTATGATACGTA 114
 |||||
 Oy 6054 GTTGTGCTGTTATATCTGTGCCCCCTTC-TTAAAGAAATGTTGATATGTTCTCTT 6112
 |||||
 Db 113 GTTGTGCTGTTATATCTGTGCCCCCTTTTAAAGAAATGTTGATATGTTCTCTT 54
 |||||

Oy 6113 GGATTAATTTGATTTGATGACAACTGATTTAAATTAACATATTG 6155
 |||||
 Db 53 GGATTAATTTGATTTGATGACAACTGATTTCAATATAAATATTG 11
 |||||

RESULT 10
 BS546856 668 bp mRNA linear EST 26-OCT-2001
 LOCUS BS546856
 DEFINITION musculus cDNA full-length enriched, 0 day neonate eyeball Mus
 mRNA for SSeCKS, mRNA sequence.
 ACCESSION BS546856
 VERSION BS546856.2 GI:16447501
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 668)
 Arakawa, T., Carrinini, P., Fukuda, S., Furumoto, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda, M., Koye, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 TITLE Unpublished (2001)
 JOURNAL On Jul 31, 2000 this sequence version replaced gi:9618284.
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenitro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.go.jp,
 URL: http://genome-gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

FEATURES

source

1. .668
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="E130314119"
 /clone_lib="RIKEN full-length enriched, 0 day neonate
 eyeball"
 /tissue_type="eyeball"
 /dev_stage="0 day neonate"
 /lab_host="DH10B"
 /notes="Site 1: SalI; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGCGCGCGCACTCGAGTTTTTTTTTTTNN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adaptor of sequence [5'
 GAGAGAGAGATTCGAGTTAAATAATATCCCTCCCCCCCC 3']. cDNA
 was cleaved with BamHI and XhoI. Vector: a modified
 pBluescript KS(+) after bulk excision from Lambda FLC 1."

BASE COUNT 199 a 146 c 229 g 93 t 1 others
 ORIGIN

Query Match 8.6%; Score 529.2; DB 10; Length 668;
 Best Local Similarity 88.8%; Pred. No. 3.8e-109;
 Matches 596; Conservative 0; Mismatches 69; Indels 6; Gaps 2;

QY 1183 ATGAGAGATCGAGCCACCAAGAGTTGTTCAGAGGTCACGTGAGCACCCTGGAGA 1242
 |||||
 Db 1 ATGAGAAGCGGAGNCCACCAAGAGTTGTTCAGAGGTCACGTGAGCACCCTGGAGA 60
 |||||
 QY 1243 AGACAGAGGAGGAGCGGAGGAGGAGGAGGCTGAAGGGGGCGTGTGTGTAGAGGAA 1302
 |||||
 Db 61 AGATGACGAAGGCA---AGAGAGCAGAGGTGGAGGGGATGTGTGTGTGAAGAT 117
 |||||

QY 1303 CAGGAGATCTTTGCCCCCTGAGAAATGGCTGAGCCCCAGGAGTCCCCCAGGAGCTG 1362
 |||||
 Db 118 CGGAGAGATCTTTGCCCCCTGAGAAATGGCTGAGCCCCAGGAGTCCCCCAGGAGCTG 177
 |||||
 QY 1363 AGCTGTCTGAGGAGCTGATCAAGACGACAGAGATGTGTCTCTGGAGGAGACACCACTC 1422
 |||||
 Db 178 AGCTGTGGAGGAGCTGATCAAGACCAAGAGATGTGTCTCTGGGGGTGACCATACTC 237
 |||||
 QY 1423 AACTGACAGACTTAAGTCTCTGAAGAGAGAGCGTGGCCAAACACCCAGAGGCAATTTCA 1482
 |||||
 Db 238 AGCTGACAGATCTTAAGTCTCTGAAGAGAGAGATGTCTACCCAAACACCCCGAAGGCATTGTCA 297
 |||||
 QY 1483 GTGAGGTGGAGATGCTGTCTCTCTCAGGAAGAATCAAGGTACAGGGAAGTCCCTTGAAGA 1542
 |||||
 Db 298 GTGAGGTGGAGATGCTGTCTCTCTCAGGAGAGATCAAGGTACAGGGAAGTCCCTTGAAGA 357
 |||||
 QY 1543 AACTTCTTCACTAGCTCAGGCTTAAAGAGAGTGTCTGGGAAGAAGCAGCAAGGGAA---AC 1599
 |||||
 Db 358 AGCTTTCACAGTTTCGGGCTTAAAGAGAGTCTCCGGAAGAAGCAGCAAGGGGAAGAGAG 417
 |||||
 QY 1600 GAGGAGTGGGGGAGAGCGAAGAGCGCTTGAGAAATACCAACACATTTACACCGCAATCCCCAG 1659
 |||||
 Db 418 GAGGAGCGGGGAGATGAAGAGCCAGGAGATACCAACACATTTCAAACCGAGTCCCCAG 477
 |||||
 QY 1660 AGAGTGTGTATGACGAGAGAGGAGAGAGCTTGGTGTCTCCCGAGGAGCGCTGAGAGAGA 1719
 |||||
 Db 478 AGAGTGTGTGACGAGAGAGGAGAGAGCTTGGTGTCTCCCGAGGAGCGCGGAGAGAGA 537
 |||||
 QY 1720 CCAGTGTCTTGGAGAAAGGGCCCTCGAAGCACCCAGGATGGGAGAGCTGAGGAAGGAA 1779
 |||||
 Db 538 TCGGTGTCTTGGAGAAAGGGCCATCGAAGCACCCAGGAGCGGAAGCTGAGGAAGGAG 597
 |||||
 QY 1780 CTACTTCCGATGGAGAGAAAGAGAGAGGATCACTTCCCTGGGCAATCTTCAAAAAGA 1839
 |||||
 Db 598 CGACTTCCGAGGAGAGAGAAAGGAGGATCACTCCCTGGGCAATCTTCAAAAAGA 657
 |||||
 QY 1840 TGGTGACACCC 1850
 |||||
 Db 658 TGGTGACACCC 668
 |||||

RESULT 11
 BQ176546/c
 LOCUS
 DEFINITION UI-M-DJ2-bvw-p-22-0-UI.s1 NIH BMAP DJ2 Mus musculus cDNA clone
 UI-M-DJ2-bvw-p-22-0-UI 3', mRNA sequence.
 ACCESSION BQ176546
 VERSION BQ176546.1 GI:20352038
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 703)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov
 Tissue Procurement: Dr. Robin Davisson
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA

sequence: 1-22, >AT rich#low_complexity (matched complement)

Seq primer: M13 FORWARD

POLYA=yes

FEATURES Location/Qualifiers

source

```

1. .703
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-DJ2-bw-p-22-0-UI"
/clone_id="N1H_BMAP_DJ2"
/tissue_type="subfornical organ and postrema"
/dev_stage="Adult"
/lab_host="MDH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: brain; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-M-DJ2 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is: GCTACATGAT, subfornical
organ and area postrema.
TAG_Lib=UI-M-DJ2
TAG_TISSUE=subfornical organ and postrema
TAG_SEQ=GCTACATGAT"
BASE COUNT      249 a      127 c      124 g      203 t
ORIGIN

```

```

Query Match      8.3%; Score 512; DB 14; Length 703;
Best Local Similarity 90.6%; Pred. No. 3.1e-105;
Matches 614; Conservative 0; Mismatches 55; Indels 9; Gaps 6;

QY 5484 CTTTATGATTTATATGATTTATGATGTCCTCCCTGACCTATTTGATA-TTTT 5542
DB 703 CTTTATGATTTATATGATTTATGATGTCCTCCCTGACCTATTTGATA-TTTT 644
QY 5543 TTTCTAAGTTTAAAGCAGATGTTTGTATGATATCAATATATTAACGGTGTGACCAT 5602
DB 643 TTTCTAAGTTTAAAGCAGATGTTTGTATGATATCAATATATTAACGGTGTGACCAT 584
QY 5603 AGCGAGCTTTGAAAAGCTCCAGCCTCACTGTAACCTGACGAAACAG---ATPACA 5658
DB 583 AGCGAGCTTTGAAAAGCTCCAGCCTCACTGTAACCTGACGAAACAGAAATPACA 524
QY 5659 TTCCCTGCAAGAGAGACAGATCTTTTAAAGTTTACTGATGCTTAAGTCTGAGGCTT 5718
DB 523 TTCCCTGCAAGAGAGATCAAGTC-TTTTAAAGTTTACTGATGCTTAAGTCTGAGGCTT 465
QY 5719 CTAGTCTCTGAAAAGTGTTGTTTCTTATGACAGCGAGCTCAGAAATPAAAAACCCCAT 5778
DB 464 GTAGTCTCTGAAAAGTGTTGTTTCTTATGACAGTGAAGCTCAGAAATPAAAAACCCCAT 405
QY 5779 TTGGAACATCCAGAGATGCCAATATTAACATGATTTTTCCTCCCTTTTGTCTATCC 5838
DB 404 TTGGAACATCCAGAGATGCCAATATTAACATGATTTTTCCTCCCTTTTGTCTATCC 345
QY 5839 AGTCCAGTTGAAAAGAGTCTCTCTGTGTCAGATTAAGCCCTGCTCTTATATATATG 5898
DB 344 AGTCCAGTTGAAAAGAGTCTCTCTGTGTCAGATTAAGCCCTGCTCTTATATATG 285
QY 5899 GACCAATGATGTCCTAAGGCGATAGATGTTTCTTAATGAGAAAGAAATCTGTTGAC 5958
DB 284 GACCAATGATGTCCTAAGGCGATAGATGTTTCTTAATGAGAAAGAAATCTGTTGAC 227
QY 5959 GTTTTGTGATTTGATCTTTCTATAGTGGACCGAAATTCATATGCAAGTGAAGTATCC 6018
DB 226 ACTTTTGTGATTTGATCTTTCTATAGTGGACCGAAATTCATATGCAAGTGAAGTATCC 167
QY 6019 TGTCTTTTACAGATGATTTTGTATGATATGAGATTTGTCTGTGTATATCTGTGCC 6078

```

```

DB 166 TGCTCTTACGATGATGATTTTGTATGATAGTACGTGAGTTGTCTGTATATCTGTGCC 107
QY 6079 CTTG-TTTAAGACATGTTGATGATTTATGCTCTTGTGATTAATTTGATTTGACACTGA 6137
DB 106 CTTCTTTTAAAGAACATGTTTACATTTATGTTCTTGTGATTAATTTGATTTGACACTGA 47
QY 6138 TTTTAATTAACATATTTG 6155
DB 46 TTTCAATPAAAAATATTTG 29

```

RESULT 12

AI225639

LOCUS

DEFINITION

AI225639

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 610)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenger, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Reising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

WGI:976164

Possible reversed clone: similarity on wrong strand

Seq primer: custom primer used

High quality sequence stop: 460.

Location/Qualifiers

1. .610

/organism="Mus musculus"

/strain="C57BL"

/db_xref="taxon:10090"

/clone="IMAGE:1907968"

/clone_id="Sugano mouse kidney mkiA"

/sex="female"

/dev_stage="adult"

/lab_host="MDH10B"

/note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII

(CACTGTG); Site_2: DraIII (CACCAGTGG); 1st strand cDNA

was primed with an oligo(dT) primer

[ATGTGGCTTTTGTGTTTGTGTTT]; double-stranded cDNA was

ligated to a DraIII adaptor (TGTGGGCTGCTG); digested

and cloned into distinct DraIII sites of the pME18S-FL3

vector (5' site CACTGTG, 3' site CACCAGTGG). XhoI should

be used to isolate the cDNA insert. Size selection was

performed by Dr. Sumio Sugano (University of Tokyo

Institute of Medical Science). Custom primers for

sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end

primer CGACCTGACGCTGACGACA."

```

BASE COUNT      172 a      163 c      185 g      89 t
ORIGIN
Query Match      8.1%; Score 501.8; DB 9; Length 610;

```

Best Local Similarity 88.9%; Pred. No. 6.2e-103;		Matches 542; Conservative 0; Mismatches 68; Indels 0; Gaps 0;	
QY	2795	CATTGCAGAGAAACTCCTGTGCTTACCCAGAGCTTACCCAGAGGTTAAAGATGCCCATGA	2854
Db	1	CACCTGCAGAGCAACTCCTGCACCTCGCCAGACTTTTACAGGGGGCAAGATGCCCATGA	60
QY	2855	CGACATGGTCACCAAGTGAAGTGGATTTACCTCAGAAAGCTGTGTGACAGCCACAGAGACTC	2914
Db	61	CGACATAGTCACCAAGTGAAGTGGATTTTACCTCAGAAAGCTGTGTGACAGCCAGAGAAACAC	120
QY	2915	AGAGGCTCTCCGTACTCAAGAAGTTTACCGAAGCATCGGGGCGAAGACAGACACAGACAT	2974
Db	121	AGAGGCGCTCGCGCTGAGAACTTACCGAAGCATCAGGGGCGAAGACAGACACAGACAT	180
QY	2975	GGTGTCCGCAAGTTTCCAGCTGACTGACTCCCCAGACACACAGAGGAAGCCACCCCACT	3034
Db	181	GGTGTCTGCAGTTTCCAGCTGCTCCGACTCCCCGACACACAGAGGAGCCACCCCACT	240
QY	3035	TCAGGAGGTAGAGTGGTGTCTAGATACAGAAAGAGAGAGCGCCAGACGAGGCCAT	3094
Db	241	TCAGGAGGTAGAGGTTGGCATGCTAGATACGGAAGACAGGAGCGCCAGACGAGGCCGT	300
QY	3095	CTTCCAGCGCTTGCACACAAAGTGAAGAGGAGTCCAGGTGCTGCTCAACCCAGACTGT	3154
Db	301	CTTCCAGCGCTTGCACACAAAGTGAAGAGGAGTCCAGGTGCTGCTCAACCCAGACTCT	360
QY	3155	GCAGAGAACCGGGTCAAAAGCACTGGAGAAGGTTGAGGAGGTAGAGGAGGACTCCGAAGT	3214
Db	361	GCAGAGAGCAGGGCCGAAAGCACTGGAGAAGTGGAGAGGGTGTAGAGAGGACTTCGAGGT	420
QY	3215	GCTGGCTTCGGAGAAAGAGAGGAGCGTTATGCCAAAGGCCCGGTGCAGGAAGCTGGAGC	3274
Db	421	GCTGGCTACCCGAGAAAGAGAGGAGTGTGTGCCGAAGGACCCCGTGCAGGAAGCTGAAAC	480
QY	3275	TGAGCATCTTGCACAGGCTCTGAGACTGACAGGCTACTCCAGAGAGCCTTGAAGTTCC	3334
Db	481	TGAGCATCTTGCACAGGCTCTGAGACTGACAGGCTACTCCAGAGAGCCTTGAAGTTCC	540
QY	3335	TGAAGTCACGCGCAGATGTAGACCATGTCCGCACGTCGCGAGGTTATCAAGCTCCAGCAGCT	3394
Db	541	TGGAGTCACAGAGGATGTAGACCGTGCCACACATGCAGGTTATCAAGCACCAGCAGCT	600
QY	3395	GATGGAACAG 3404	
Db	601	GATGGAACAG 610	

RESULT 13
BM228830/c
LOCUS
DEFINITION
K0268G02-3 NIA Mouse Unfertilized Egg cDNA Library (Long) Mus
ACCESSION
BM228830
VERSION
BM228830.1
KEYWORDS
EST.
SOURCE
GI:17791601
house mouse.
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 665)
Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Stagg,C.A.,
Martin,P., Aiba,K., Tanaka,T. and Ko,M.S.H.
Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library
(Long)
JOURNAL
Unpublished (2001)
COMMENT
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0268 row: G column: 02
Seq primer: -21M13 Forward

High quality sequence stop: 665 POLYA=Yes.		Location/Qualifiers	
FEATURES		source	
1..665		/organism="Mus musculus"	
		/strain="C57BL/6J"	
		/db_xref="niaEST:K0268G02-3"	
		/db_ref="taxon:10090"	
		/clone="K0268G02"	
		/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long)"	
		/tissue_type="Unfertilized Egg"	
		/lab_host="DH10B"	
		/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun-grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]. 5'-pGACTAGTTCTAGATCGAGCGCGCCCTTTTCTTTT-3'), treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."	
BASE COUNT		237 a	126 c 120 g 182 t
ORIGIN			
Query Match		8.0%;	Score 492.2; DB 13; Length 665;
Best Local Similarity		90.3%;	Pred. No. 9.4e-101;
Matches 606;		Conservative	0; Mismatches 53; Indels 12; Gaps 7;
QY	5489	TGTAATTTATGATGTTTTATGATGCTCTCTCTGTTACCTATTGTTATA-TTTTTTCT	5547
Db	665	TGTAATTTATGATGTTTTATGATGCTCTCTCTGTTACCTATTGTTATA-TTTTTTCT	606
QY	5548	AACGTTTAAAGCACATGCTTTTGTATATGCAATATATACGGGTGTGAGCCATAGCGA	5607
Db	605	AAGGTTTAAAGCACATGCCCTTTTGTATATGCAATATATACGGGTGTGAGCCATAGCGA	546
QY	5608	CGCTTTGAAAAGCTCCCAAGCCTCAACTGTAACTGCAGCAACAG-ATACATTCCT	5663
Db	545	AGCTTTGAGAAGCCCCAAGCCTCAACTGTAACTGCAGCAACAGCAATAACATTCCT	486
QY	5664	GGCAAGAAGACACAAGTCTTTTAAAGTTTACTGATGCTTAGATCTGTGGGCTCTAGT	5723
Db	485	GGCAGGAAGATACAACTC-TTTTAAAGTTTACTGATGCTTAGATCTGTGGGCTCTAGT	427
QY	5724	CCTCTGAAAAGTGGTGGTTTTCCTATGCAAGCGAGCTCAGAAATAAAAAACCCCATTTGA	5783
Db	426	CCTCTGAAAAGTGGTGGTTTTCCTATGCAAGCTGAGCTCAGAAATAAAAAACCCCATTTGA	367
QY	5784	ACATCCAGGATGTCCCAATATATACCATGATTTTTCCTCCCTTTTGGTCTAATCCAGTCC	5843
Db	366	GACATTCGAATGTCCCAACATTTACGACAACATTTTTT---TTTTTTTCTTAATCCAGTCC	310
QY	5844	AGGTTGGAAGAGTCTCCTCTGTGTGATGATTAAGCCCTGTCTCTTAATGATATGACAA	5903
Db	309	AGGTTGGAAGAGTCTCCTCTGTGTGATGATTAAGCCCTGTCTCTTAATGATATGACAG	250
QY	5904	ATGAGTGTGCTTAAGGCCATGAGATGTTTCTTAATGAGAAGGAATCTGTGTGACGTTTT	5963

Db	249	ATGAGTGGCC-ATGGCCATGAGCTATTTCCTAAAGCAGAGGAATTTGTGT-TACTTT	192
Qy	5364	TTTGGTTTACTCTTCTCTATGCTGGACCGAATTCATATGACATGCAAGGAGATCCTGTTTC	6023
Db	191	TTGGATTGTACTCTTCTTAGGCTGGACCGAATTCATATGACATGCAAGGAGATCCTGCTC	132
Qy	6024	TTTACAGATGGATTTTGATATAGATCTGAGATTGTCTGTATATCTGTGCCCTTC-	6082
Db	131	TTTACAGATGGATTTTGATATAGATCTGAGATTGTCTGTATATATCTGTGCCCTTC	72
Qy	6083	TTTAAGAACATGTGTCATTATGTCTCTTTGGATTAATGTGTGATTTGACAACTGATTTAA	6142
Db	71	TTTAAGAACAAATGTTACATTATGTCTCTTTGGATTAATGTGTGATTTGACAACTGATTTCA	12
Qy	6143	ATTAACATATT 6153	
Db	11	ATTAATAATTTT 1	

RESULT 14					
AI103489/c	AI103489	476 bp	mRNA	linear	EST 08-JAN-1993
LOCUS	EST212778	Normalized rat embryo,	Bento Soares	Rattus sp.	cDNA clone
DEFINITION	REMCC63 3' end, mRNA sequence.				
ACCESSION	AI103489				
VERSION	AI103489.1	GI:4134176			
KEYWORDS	EST.				
SOURCE	Rattus sp.				
ORGANISM	Rattus sp.				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 476)
Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat
Gene Index
Unpublished (1998)
On Oct 6, 1998 this sequence version replaced gi:3708028.

JOURNAL
COMMENT

Unpublished (1998)
On Oct 6, 1998 this sequence version replaced gi:3708028.
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.

FEATURES	source	Location/Qualifiers
		1..476
		/organism="Rattus sp."
		/db_xref="taxon:10118"
		/clone="REMCC63"
		/clone_id="Normalized rat embryo, Bento Soares"
		/dev_stage="embryo 8, 12, 18 dpc"
		/note="Vector: pPT73Pac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT		170 a 94 c 90 g 122 t
ORIGIN		

Query Match	7.66	Score 467.8	DB 9	Length 476
Best Local Similarity	99.66	Pred. No. 3.1e-95		
Matches 469	Conservative 0	Mismatches 2	Indels 0	Gaps
Qy	5688	AAAGTTACTGATGCTGATCTGTGGGCTTCTAGTCTCTGAAAGTGTTGTTTCTTA	5747	
Db	476	AAAGTTACTATGTTTATGATCTGTGGGCTTCTAGTCTCTGAAAGTGTTGTTTCTTA	417	
Qy	5748	TGCACGCGAGCTCAGAAATATAAAAACCCATTTTGAAACATCCAGATGTCCAAATATTA	5807	
Db	416	TGCACGCGAGCTCAGAAATATAAAAACCCATTTTGAAACATCCAGATGTCCAAATATTA	357	
Qy	5808	CCATGATTTTTTCCCCCTTTTGTGATCCAGTCCGAGTTGGAAAAAGATCTCTCTGT	5867	
Db	356	CCATGATTTTTTCCCCCTTTTGTGATCCAGTCCGAGTTGGAAAAAGATCTCTCTGT	297	

QY	5868	GTGGATTAAGCCCTCTCTCTTAAAGATGTGACAAAGAGTGATCCTAAGGCCATGAGA	5927
Db	296	GTGAGATTAAAGCCCTCTCTCTTAAAGATGTGACAAAGAGTGATCCTAAGGCCATGAGA	237
QY	5528	TGTTTCTTAATGACAAAGAAATCTGTGTGACGTTTTTTTGATTGTACTCTCTATGCTGG	5987
Db	236	TGTTTCTTAATGACAAAGAAATCTGTGTGACGTTTTTTTGATTGTACTCTCTATGCTGG	177
QY	5988	ACCGAATTCATGACGAGATCGAAGTAGAGCTGTCCTTACAGATGATTTTGATAGAT	6047
Db	176	ACCGAATTCATGACGAGATCGAAGTAGAGCTGTCCTTACAGATGATTTTGATAGAT	117
QY	6048	ACTGAGATTGTCTGTGTATATCTGTGCCCTTCTTTAAGAACATGTTGATTATGTT	6107
Db	116	ACTGAGATTGTCTGTGTATATCTGTGCCCTTCTTTAAGAACATGTTGATTATGTT	57
QY	6108	CCTTGGATTAATTTGATTTGCAACTGATTTTAAATAAATATTTTACT	6158
Db	56	CCTTGGATTAATTTGATTTGCAACTGATTTTAAATAAATATTTTACT	6

RESULT	15
BG670384	
LOCUS	BG670384
DEFINITION	BG670384 573 bp mRNA linear EST 30-APR-2001 DRNBAB11 Rat DRG Library Rattus norvegicus cDNA clone DRNBAB11 5' , mRNA sequence.
ACCESSION	BG670384
VERSION	BG670384.1 GI:13892483
KEYWORDS	EST .
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus

REFERENCE	1 (bases 1 to 573)
AUTHORS	Xiao, H.S., Huang, Q.H., Zhang, F.X., Bao, L., Lu, Y.J., Guo, C., Yang, L.,
TITLE	Identification of gene expression profile of dorsal root ganglion in the rat peripheral axotomy model of neuropathic pain
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (12), 8360-8366 (2002)
MEDLINE	22056133
COMMENT	Contact: Zhang Xu

COMMENT

Contact: Zhang Xu
Laboratory of Sensory System
Institute of Neuroscience
320 Yue Yang Road, Shanghai 200031, P.R.China
Tel: 86-21-64748700-121
Fax: 86-21-64713446
Email: xu.zhang@ion.ac.cn

This clone is also available at Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong New Area, P.R. China. Please contact with zhang Xu (xu.zhang@ion.ac.cn) or Han Zeguang (hanzg@chc.sh.cn)

PCR Primers
FORWARD: T3
BACKWARD: T7
Seq primer: T3
POLYA-No.

```

FEATURES
source
location/Qualifiers
1..573
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="DRNBAB11"
/clone_lib="Rat DRG Library"
/sex="male"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"
/note="Total RNA was isolated from hypothalamus and
transcribed into cDNA, which was then used as template in
PCR. The PCR products were separated on sequencing gel.
The differential bands were cut, reamplified, cloned into
pMD18-T vector and confirmed by Northern blot."

```

BASE COUNT 151 a 131 c 115 g 170 t 6 others
ORIGIN

Query Match 7.5%; Score 460.2; DB 12; Length 573;
Best Local Similarity 88.0%; Pred. No. 1.7e-93;
Matches 557; Conservative 0; Mismatches 14; Indels 62; Gaps 3;

Qy 5088 GAATCCTAAGATGTAGTGTCTGCTGATTTGATCTGTGAAGACCAAGTGTGAACAAGTC 5147
Db 1 GAATCCTAAGATGTAGTGTCTGCTGATTTGATCTGTGAAGACCAAGTGTGAACAAGTC 60

Qy 5148 ACAGAACAAAGATGCTGCTGTTGGGACCTTGAGACCAAGATTTGAGGCCATGAGATCCA 5207
Db 61 ACAGAACAAAGATGCTGCTGTTGGGACCTTGAGACCAAGATTTGAGGCCATGAGATCCA 120

Qy 5208 GAGAGCGGGCCGTCCTAATGATTTCCACCCAGTAGACACCCCGACAAATTCAGAGGCTTC 5267
Db 121 GAGAGCGGGCCGTCCTAATGATTTCCACCCAGTAGACACCCCGACAAATTCAGAGGCTTC 180

Qy 5268 ATCGGGAGCTAGACCCAGCTTAACATTTCTCGTTTCAAGACTGCTTTGATTTGCCCTT 5327
Db 181 ATCGGGAGCTAGACCCAG----- 198

Qy 5328 GATCCCGTCCGTGTATTTCTAACAATTTCTCGTTTCAAGACTGCTTTGATTTGCCCTT 5387
Db 199 -----CTAACAATTTCTCGTTTCAAGACTGCTTTGATTTGCCCTT 240

Qy 5388 GATCCCGTCCGTGTATTTCCGATTTAAGTCTCGTTTCTCAACCTGGAACCAAT-TCTG 5446
Db 241 GATCCCGTCCGTGTATTTCCGATTTAAGTCTCGTTTCTCAACCTGGAACCAATGCTG 300

Qy 5447 CCATACCTAGTTCACATCTCAAACTGGAGCATCCTCTTTATGTATTTATGTATGTT 5506
Db 301 CCATACCTAGTTCACATCTCTCAAACTGGAGCATCCTCTTTATGTATTTATGTATGTT 360

Qy 5507 TTATGTAGTCTCTCTGTAACCTATTGTATTTTCTTAACGTTTAAACCATGCTT 5566
Db 361 NTATGTAGNCCCTCTGTAACCTATTGTATTTTCTTAACGTTTAAACCATGCTT 420

Qy 5567 TTTGTATTATGCAATATATAACGGGTGTGAGCCATAGCGACGCTTTGAAAAGCTCCAAG 5626
Db 421 TTTGTATTATGCAATATATAACGGGTGTGAGCCATAGCGACGCTTTGAAAAGCTNCAAG 480

Qy 5627 CCTCAACTGTAACTGAGCAAAACAGATAACA-TTCCTGGCAAGAGAGACAAGCTTTTT 5685
Db 481 CCTCAACTGTAACTGAGCAAAACAGATAACA-TTCCTGGCAAGAGAGACAAGCTTTTT 540

Qy 5686 TTAAGGTTTACTGATGCTTAGATCTGTGGCTT 5718
Db 541 TTAAGGNTACTGATGCTTAGATCTGTGGCTT 573

Search completed: December 12, 2002, 16:19:41
Job time : 7377 secs

THIS PAGE BLANK (uspro)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2002, 16:39:22 ; Search time 123 Seconds
(without alignments)
1729.008 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAQSSTEQSPQAGSDT.....AWAQRKCLRLQLKAPVSK 1596

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8049	99.7	1596	AAW31347	Rat tumour suppress
2	5387.5	66.7	1346	AAW31346	Rat tumour suppress
3	4334	53.7	1783	ABR97448	Novel human protei
4	4294	53.2	1781	AAQ17365	Human gravin. Hom
5	4274.5	52.9	1780	AAW53863	Human gravin polyp
6	4274.5	52.9	1780	AAW53863	Human gravin prote
7	4262.5	52.8	1795	ABG21018	Novel human diagno
8	952.5	11.8	292	AAW53862	Gravin polypeptide
9	952.5	11.8	292	AAW53779	Recombinant human
10	656	8.1	651	AAW53875	Gravin polypeptide

11	656	8.1	651	21	ABW18410	Human gravin PKA r
12	642	8.0	2768	19	ABW68397	Drosophila melanog
13	639.5	7.9	453	19	AAW53867	Gravin polypeptide
14	639.5	7.9	453	21	ABW18407	Human gravin PKA r
15	557.5	6.9	6815	22	ABW66811	Drosophila melanog
16	521.5	6.5	396	19	AAW53868	Gravin polypeptide
17	521.5	6.5	396	21	ABW18408	Human gravin PKA r
18	492	6.1	2519	22	ABG16636	Novel human diagno
19	485	6.0	3263	22	ABW67210	Drosophila melanog
20	462	5.7	1786	18	AAW24790	P. falciparum live
21	462	5.7	1787	23	AAU96699	Plasmodium falcipo
22	451	5.6	1468	22	ABW62991	Drosophila melanog
23	440.5	5.5	1812	22	ABW58022	Drosophila melanog
24	431	5.3	3257	22	ABW67502	Drosophila melanog
25	430	5.3	3111	22	ABW60327	Drosophila melanog
26	429	5.3	2665	22	ABW33490	Human peptide #965
27	429	5.3	2665	22	ABW33490	Human peptide #965
28	429	5.3	2665	22	ABW18950	Protein #949 encod
29	429	5.3	2665	22	AAW54270	Human brain expres
30	429	5.3	2665	22	AAW66665	Human bone marrow
31	429	5.3	2665	22	AAW14533	Peptide #967 encod
32	429	5.3	2665	22	AAW26950	Peptide #987 encod
33	429	5.3	2665	22	AAW02259	Peptide #941 encod
34	429	5.3	2665	23	ABG36319	Human peptide enco
35	429	5.3	2748	22	ABW58843	Drosophila melanog
36	429	5.3	3266	21	AAW42491	Human ORFX ORF2255
37	421	5.2	1714	22	ABG02138	Novel human diagno
38	421	5.2	1714	22	ABG06417	Novel human diagno
39	420	5.2	2803	23	ABW08161	Human cytoskeleton
40	414.5	5.1	2622	22	ABG06418	Novel human diagno
41	411	5.1	2618	22	ABG02135	Novel human diagno
42	409.5	5.1	2515	22	ABW12281	Human secreted pro
43	409.5	5.1	2515	22	AAW80268	Human protein SEQ
44	409.5	5.1	2515	22	AAW80268	Human protein SEQ
45	409	5.1	1793	22	ABW59613	Drosophila melanog

ALIGNMENTS

RESULT 1	AAW31347	AAW31347 standard; Protein; 1596 AA.
ID	AAW31347	
XX	AAW31347;	
XX	AC	
XX	20-JUL-1998 (first entry)	
DT		
XX	Rat tumour suppressor protein SseCKs.	
DE		
XX	SseCKs; tumour suppressor gene; rat; protein kinase C; mitosis; cancer; malignancy; cell proliferation; Alzheimer's disease; therapy.	
KW		
XX	Rattus sp.	
OS		
XX		
FT	Key	Location/Qualifiers
FT	Peptide	1..10
FT	Region	/notes="myristylation site"
FT		410..418
FT	Peptide	/note="glycine-rich region"
FT		517..520
FT		/notes="nuclear localisation signal"
FT	Region	524..527
FT		/notes="glycine-rich region"
FT	Peptide	677..680
FT		/notes="nuclear localisation signal"
FT	Peptide	757..760
FT		/notes="nuclear localisation signal"
FT	Region	1399..1421
FT		/note="zinc finger"
FT	Modified-site	279..307
FT		/notes="proposed protein kinase C phosphorylation"

FT Modified-site 504..526 site"
FT /note="proposed protein kinase C phosphorylation
FT site"
FT Modified-site 592..614
FT /note="proposed protein kinase C phosphorylation
FT site"
FT Modified-site 741..766
FT /note="proposed protein kinase C phosphorylation
FT site"
FT MO9740059-A1.
FT 30-OCT-1997.
FT 18-APR-1997; 97WO-US06830.
FT 18-JUN-1996; 96US-0665401.
FT 19-APR-1996; 96US-0635121.
XX (GELM/) GELMAN I.
XX (JAKE/) JAKEN S.
XX Gelman I, Jaken S;
XX WPI: 1997-535770/49.
XX N-PSDB: AAV02302.
XX Tumour suppressor gene SSeCKS - used as a mitotic regulator, and
XX inhibitor of malignant phenotype
XX Claim 4; Fig 11A-L; 162pp; English.
XX This polypeptide comprises the novel tumour suppressor protein
XX SSeCKS that is a substrate of protein kinase C and which acts as a
XX negative regulatory of mitosis and as an inhibitor of the
XX transformed phenotype. Its amino acid sequence was deduced from an
XX isolated nucleic acid molecule (see AAV02302). The SSeCKS nucleic
XX acid, as well as homologous and hybridising nucleic acids are
XX claimed, as are isolated proteins encoded by such nucleic acids,
XX vectors comprising the nucleic acids, host cells, and methods of
XX inhibiting the expression of a transformed phenotype in a host cell
XX by introducing the nucleic acids. Introduction of a SSeCKS nucleic
XX acid or gene product into a host cell inhibits mitosis of the host
XX cell, allowing the treatment of diseases associated with disorders
XX of proliferation and/or with the expression of a malignant
XX phenotype. SSeCKS can also be used to treat or identify disorders
XX of cytoskeletal structure and cellular architecture (such as
XX Alzheimer's disease), and may be a marker for aberrancies in
XX fertility and/or nervous system development.
SQ Sequence 1596 AA;
Query Match 99.7%; Score 8049; DB 18; Length 1596;
Beet Local Similarity 99.8%; Pred. No. 0;
Matches 1593; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 241 LQASDDAAAEBAKDEGEKEKEPTKSPSPSPSPVNSSETTSSPKKFFPTGMAGMRKKT 300
Db 241 LQASDDAAAEBAKDEGEKEKEPTKSPSPSPSPVNSSETTSSPKKFFPTGMAGMRKKT 300
Qy 301 FKSKEDDLETAERKQKQAEKVDKEEKEKTEPASEQEPADTDQARLSADYKEVELPL 360
Db 301 FKSKEDDLETAERKQKQAEKVDKEEKEKTEPASEQEPADTDQARLSADYKEVELPL 360
Qy 361 EDQVGDLEASSEKCAPLATEVPDEKMEAHQEVVAEYHVHVTVEKTEBEQGGGABEGV 420
Db 361 EDQVGDLEASSEKCAPLATEVPDEKMEAHQEVVAEYHVHVTVEKTEBEQGGGABEGV 420
Qy 421 VEGTGESLPEPKLAEPOEVPOBAEPARELMKSRMVCVSGGDHTQTLSPBEKTLPHGPE 480
Db 421 VEGTGESLPEPKLAEPOEVPOBAEPARELMKSRMVCVSGGDHTQTLSPBEKTLPHGPE 480
Qy 481 GIVESEVMSQERIKVQSGPLKKLFSSSGIKLKGKQKQKRGCGGDEBEGEQHITHE 540
Db 481 GIVESEVMSQERIKVQSGPLKKLFSSSGIKLKGKQKQKRGCGGDEBEGEQHITHE 540
Qy 541 SPESADEQKGESSASPEBPEETTCLKGLPEAPQDGEABEGTTSDEKKEGITPMASF 600
Db 541 SPESADEQKGESSASPEBPEETTCLKGLPEAPQDGEABEGTTSDEKKEGITPMASF 600
Qy 601 KKMVTPKKRVARPSBSDKKEELEKVSATLSTDTVSEMQDEKTVGEBQKPREPKRV 660
Db 601 KKMVTPKKRVARPSBSDKKEELEKVSATLSTDTVSEMQDEKTVGEBQKPREPKRV 660
Qy 661 DTSVSEALICVGSKKRARKKASSDDEGPRTLGDSHRAEASKQKEAGTDVAPASTQ 720
Db 661 DTSVSEALICVGSKKRARKKASSDDEGPRTLGDSHRAEASKQKEAGTDVAPASTQ 720
Qy 721 EODQAQSSSPPEPAGSPSEBEGCVSTWESFKRLVTPRKKSKSLEKAKDDSSVEQLSTEIE 780
Db 721 EODQAQSSSPPEPAGSPSEBEGCVSTWESFKRLVTPRKKSKSLEKAKDDSSVEQLSTEIE 780
Qy 781 PSRESWVSIKKFLPGRKKRKGADQEOATVEDSGPVEINDDPNVAVPLSEYNVER 840
Db 781 PSRESWVSIKKFLPGRKKRKGADQEOATVEDSGPVEINDDPNVAVPLSEYNVER 840
Qy 841 EKMEAGQNTLPLQLGAVVYSEELSKTLVHTVSAVAVDGTAVTSVERSPWISASYTE 900
Db 841 EKMEAGQNTLPLQLGAVVYSEELSKTLVHTVSAVAVDGTAVTSVERSPWISASYTE 900
Qy 901 PLEHTAGAMPVBEVTEKDIIAETEPVLTQTLPEGDAHDDMTSEVDFSEAVTATET 960
Db 901 PLEHTAGAMPVBEVTEKDIIAETEPVLTQTLPEGDAHDDMTSEVDFSEAVTATET 960
Qy 961 SEALRTEEVTEASGAEEETTDVNSAVSQTLDSPDTTEATPVQEVESGVLDTSEEBRQQA 1020
Db 961 SEALRTEEVTEASGAEEETTDVNSAVSQTLDSPDTTEATPVQEVESGVLDTSEEBRQQA 1020
Qy 1021 ILQAVADVKEESQVPATQVQRTGSKALEKVEVEEEDSEVLASEKEKDVWPKPVQAG 1080
Db 1021 ILQAVADVKEESQVPATQVQRTGSKALEKVEVEEEDSEVLASEKEKDVWPKPVQAG 1080
Qy 1081 AEHLAQSGSETQATPESLEVEPVDADVHVATCOVIKIQQLMEQAVAPESSETLTDSETN 1140
Db 1081 AEHLAQSGSETQATPESLEVEPVDADVHVATCOVIKIQQLMEQAVAPESSETLTDSETN 1140
Qy 1141 GSTPLADSDTADGTOQDDTTDSQDSKATAAAROSQVTEEBATQKEBPSTLPNNVPAQE 1200
Db 1141 GSTPLADSDTADGTOQDDTTDSQDSKATAAAROSQVTEEBATQKEBPSTLPNNVPAQE 1200
Qy 1201 EHGEPRGADVLEPTQOELTTAAVAVLAKTEVQGEVGDVMDLGEKVEQAEVFGVSGPNSQ 1260
Db 1201 EHGEPRGADVLEPTQOELTTAAVAVLAKTEVQGEVGDVMDLGEKVEQAEVFGVSGPNSQ 1260
Qy 1261 KAADVTYDSEVWVAGCOKEKSTEVQSLSEEGEMETDVEKEKETPKQVSEBGEQETA 1320
Db 1261 KAADVTYDSEVWVAGCOKEKSTEVQSLSEEGEMETDVEKEKETPKQVSEBGEQETA 1320
Qy 1321 APHEGTYGKVLTLDMPSERKALGSLGSPSLPDDKAGCIEVQVQSILDTVTQTAE 1380

```

Db 1321 APEHEGTYGKPVLTLDMESSRGKALGSLGGSPSLPDDQKAGCIEVQVSQSLDTTQTAE 1380
Qy 1381 AVEKVIETWISGETSEPCVGAHLLPAEKSSATGGHWTLOHAEDTVPGLPESQAESPI 1440
Db 1381 AVEKVIETWISGETSEPCVGDHLLPAEKSSATGGHWTLOHAEDTVPGLPESQAESPI 1440
Qy 1441 IVTPAPESTLHPDLQGEISASQSRSEBEDPDAQPDADGKESIAIEKVLKAEPEILELE 1500
Db 1441 IVTPAPESTLHPDLQGEISASQSRSEBEDPDAQPDADGKESIAIEKVLKAEPEILELE 1500
Qy 1501 SKSNKIVLNLQTAVDQPARTETAPETHAYDSQTPVACRLDSREPNCWTMKMDAKMKH 1560
Db 1501 SKSNKIVLNLQTAVDQPARTETAPETHAYDSQTPVACRLDSREPNCWTMKMDAKMKH 1560
Qy 1561 PVPQPREDLQVLTLEAWAQPCKLPRLQLKAPVSK 1596
Db 1561 PVPQPREDLQVLTLEAWAQPCKLPRLQLKAPVSK 1596

RESULT 2
AAW31346
ID AAW31346 standard; Protein; 1346 AA.
XX
AC AAW31346;
XX
XX
DT 20-JUL-1998 (first entry)
DE Rat tumour suppressor protein SSeCKs (active truncated form).
XX
XX SSeCKs; tumour suppressor gene; rat; protein kinase C; mitosis;
KW cancer; malignancy; cell proliferation; Alzheimer's disease;
KW therapy.
XX
XX Rattus sp.
XX

Key Location/Qualifiers
FT Misc-difference 72 /note= "encoded by TCT"
FT Misc-difference 496 /note= "encoded by GTC"
FT Misc-difference 499 /note= "encoded by GTC"
FT Misc-difference 757 /note= "encoded by CCC"
FT Misc-difference 785 /note= "encoded by AGG"
FT Region 24..32 /note= "glycine-rich region"
FT Peptide 131..134 /note= "nuclear localisation signal"
FT Region 138..141 /note= "glycine-rich region"
FT Peptide 202..205 /note= "nuclear localisation signal"
FT Peptide 289..292 /note= "nuclear localisation signal"
FT Region 306..315 /note= "glycine-rich region"
FT Peptide 368..371 /note= "nuclear localisation signal"
FT Region 1013..1035 /note= "zinc finger"
XX WO9740059-A1.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-US06830.
XX
XX 18-JUN-1996; 96US-0665401.
XX
XX 19-APR-1996; 96US-0635121.
XX

```

```

PA (GELM/) GELMAN I.
PA (JAKE/) JAKEN S.
XX
XX Gelman I, Jaken S;
XX
XX WPI; 1997-535770/49.
XX N-PSDB; AAV02301.
XX
XX Tumour suppressor gene SSeCKs - used as a mitotic regulator, and
XX inhibitor of malignant phenotype
XX
XX Example 6; Fig 3A-G; 162pp; English.
XX
XX This polypeptide comprises an active truncated form of the novel
XX tumour suppressor protein SSeCKs that is a substrate of protein
XX kinase C and which acts as a negative regulatory of mitosis and as
XX an inhibitor of the transformed phenotype. Its amino acid sequence
XX was deduced from a partial cDNA clone (see AAV02301) obtained from
XX NIH3T3 cells. The full-length SSeCKs sequence (see AAW31347) is also
XX provided. Full-length SSeCKs nucleic acid (see AAV02301), as well as
XX homologous and hybridising nucleic acids are claimed, as are
XX the nucleic acids, host cells, and methods of inhibiting the
XX expression of a transformed phenotype in a host cell by introducing
XX the nucleic acids. Introduction of a SSeCKs nucleic acid or gene
XX product into a host cell inhibits mitosis of the host cell,
XX allowing the treatment of diseases associated with disorders of
XX proliferation and/or with the expression of a malignant phenotype.
XX SSeCKs can also be used to treat or identify disorders of
XX cytoskeletal structure and cellular architecture (such as
XX Alzheimer's disease), and may be a marker for aberrancies in
XX fertility and/or nervous system development.
XX
XX Sequence 1346 AA;
XX
Query Match 66.7%; Score 5387.5; DB 18; Length 1346;
Best Local Similarity 91.1%; Pred. No. 1.8e-279;
Matches 1092; Conservative 25; Mismatches 67; Indels 15; Gaps 7;
Qy 387 MEAHOEVVAEVHVSTVEKTEEBEGGGEAGGVVVEGTGSLPPEKLAEPQVEPQEAEP 446
Db 1 MEAHOEVVAEVHVSTVEKTEEBEGGGEAGGVVVEGTGSLPPEKLAEPQVEPQEAEP 60
Qy 447 EELMKSRMCVSGGDHTQLTDLSPPEKTLPKHPGIVSEVEMLSQSRIRIKVQSPLKLF 506
Db 61 EELMKSRMCVSGGDHTQLTDLSPPEKTLPKHPGIVSEVEMLSQSRIRIKVQSPLKLF 120
Qy 507 SSSGLKKLGGKKGKGGGDEEPGEYQIHITESPESADEQKGESASSPEEPTTCL 566
Db 121 SSSGLKKLGGKKGKGGGDEEPGEYQIHITESPESADEQKGESASSPEEPTTCL 180
Qy 567 EKGPLEAPODGEABEGTTSDEKKRE---GITPWASFKKMVTPKKVRPSPSSDKEEL 622
Db 181 EKGPLEAPRMGKLRKELLR-GEKKRKDHSLGI-----LQKGDGTQETVRPSSDKEEL 234
Qy 623 EKVKSATLSSTDSTVSEMDEVKTVGEEQKPEEPKRRVDTSVSWEALICVSSSKRKARKA 682
Db 235 EKVKSATLSSTDSTVSEMDEVKTVGEEQKPEEPKRRVDTSVSWEALICVSSSKRKARKA 294
Qy 683 SSSDDEGGPRTLGGDSHRAEASKDXEAGTDVAPASTQEQDAQSSSPSPAGSPSEGE 742
Db 295 SSSDIR-GRPRLGGGQSQRGQQRQRRTDAVPASTQEQDAQSSSPSPAGSPSEGE 353
Qy 743 VSTWESFKRLVTPRKSKSKLEEK-AEDSSVEQLSTIEPESREESWVS--IKKPIGRRK 799
Db 354 VSTWESFKRLVTPRKSKSKLEEKAGRTLIVGAGCPLRSNRVNEKNLGFRLNRSPPDGR 413
Qy 800 KRADGKQEQATVEDSGPVEINEDDPNVPAVPLSEYNAVEREKWMAQNTLPOLLCAVY 859
Db 414 KQMGREQEQATVEDSGPVEINEDEPDVPAVPLSEYDAVEREKWMAQNELSCWGCV- 472
Qy 860 VSELSKTLVHTVSVAVIDCTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEK 919

```

Db 473 VSEBISKLTVHTVSAVIDGRAKTSKEERSPSWISASVTEPLEHTAGAMPVEVTEK 532
Qy 920 DIIAETPLVLTQLEPGDAHDMVTSEVDTSEAVTAETSEALRTEVTEASAEETT 979
Db 533 DIIAETPLVLTQLEPGDAHDMVTSEVDTSEAVTAETSEALRTEVTEASAEETT 592
Qy 980 DMVSAVSGLTSPDTTEATPVQEVESGVLDTREEROTOMILOAVAKVVEESQVPNTQ 1039
Db 593 DMVSAVSGLTSPDTTEATPVQEVESGVLDTREEROTOMILOAVAKVVEESQVPNTQ 652
Qy 1040 TVORTGSALKEVEEVEDSEVLASEKEDVMPKGPVQEGAEHLAQSEGTQATPEELE 1099
Db 653 TVORTGSALKEVEEVEDSEVLASEKEDVMPKGPVQEGAEHLAQSEGTQATPEELE 712
Qy 1100 VPEVTADVAVATCCVILKIQOLMEQAVAPESSETLTJSETNGSTPLADSDTADGTQDET 1159
Db 713 VPEVTADVAVATCCVILKIQOLMEQAVAPESSETLTJSETNGSTPLADSDTADGTQDET 772
Qy 1160 IDSQDSKATAAVRQSVTEBEAATAQKEPSTLPNNVAQEHGEEPRDYLEPTQOELT 1219
Db 773 IDSQDSKATAAVRQSVTEBEAATAQKEPSTLPNNVAQEHGEEPRDYLEPTQOELT 832
Qy 1220 AAADVVLAKTEVGEGEVDMLDGEKVEKEQEVFVHSGPNSOKADVTYDSEVMGVAQCOE 1279
Db 833 AAADVVLAKTEVGEGEVDMLDGEKVEKEQEVFVHSGPNSOKADVTYDSEVMGVAQCOE 892
Qy 1280 KESTEVQSLSEEGEMETDVEKEKRETKPEQVSEBEOETAPEHEGTGYKPEVLTLDMP 1339
Db 893 KESTEVQSLSEEGEMETDVEKEKRETKPEQVSEBEOETAPEHEGTGYKPEVLTLDMP 952
Qy 1340 SERGALSLGSPSLPPRODKAGCTEVQVQSLDTVTQTAENAVKIVLETVISSEGESPE 1399
Db 953 SERGALSLGSPSLPPRODKAGCTEVQVQSLDTVTQTAENAVKIVLETVISSEGESPE 1012
Qy 1400 CVGAHLPAEKSSATGCHWTLQHAEDTVPLGPESQAESIPITVPAPSTLHPDLQGETS 1459
Db 1013 CVGAHLPAEKSSATGCHWTLQHAEDTVPLGPESQAESIPITVPAPSTLHPDLQGETS 1072
Qy 1460 ASQRESEEDKPDAGPADGKSTALEKVLKAEPELLELESKSNKIYLVNTQTAVDQFA 1519
Db 1073 ASQRESEEDKPDAGPADGKSTALEKVLKAEPELLELESKSNKIYLVNTQTAVDQFA 1132
Qy 1520 RTEHPETHAYNSQTPACRLDSREPNRCWTMKMDAKMKHVPQPRBDLYLTLEAM 1578
Db 1133 RTEHPETHAYNSQTPACRLDSREPNRCWTMKMDAKMKHVPQPRBDLYLTLEAM 1191

RESULT 3

ID ABB97448 standard; Protein; 1783 AA.

AC ABB97448;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 716.

Human; anti-inflammatory; anti-inflammatory; immunomodulator;
anti-inflammatory; cerebroprotective; cytoprotective; rheumatic; gene therapy;
neuroprotective; antiparkinsonian; protein therapy; EST;
expressed sequence tag.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-292408/33.
DR N-PSDB; ABB97448.
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX Claim 20; SEQ ID NO 716; 509pp; English.
PS The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention.
CC
XX
SQ Sequence 1783 AA;
Query Match 53.7%; Score 4334; DB 23; Length 1783;
Best Local Similarity 58.2%; Pred. No. 5,6e-223;
Matches 962; Conservative 188; Mismatches 401; Indels 102; Gaps 32;
Qy 1 MGAGSTGORSPEQ-PAGSDPSEVLVSGHGAFAEA-GAAGDPADA--DPATLPKNG 56
Db 1 MGAGSTGORSPEQ-PAGSDPSEVLVSGHGAFAEA-GAAGDPADA--DPATLPKNG 60
Qy 57 QLSVNGVAGQDVHVOENOEQ-----EEVDEVDYQRESEVDREKDR 102
Db 61 QLSVNGVAGQDVHVOENOEQ-----EEVDEVDYQRESEVDREKDR 120
Qy 103 VEEAANSTAVEDITKQGEETSEIIQIPASENNVEMQVPAQNDVGFKKVFRVVG 162
Db 121 DKEAATSAVAVHDTTDDQEEPTPEIIQIPASENNVEMQVPAQNDVGFKKVFRVVG 180
Qy 163 FKFTVKDKNEKSDPTVOLLTVKQDEGAEASVAGAHQSPSVTAAGSASKESEKOS 222
Db 181 FKFTVKDKNEKSDPTVOLLTVKQDEGAEASVAGAHQSPSVTAAGSASKESEKOS 235
Qy 223 TEKQEGTLKQESSTEIPLQAESDQAAEEBAKQDEGEKEKEPTSPSPSSPVNSETTS 282
Db 236 TEKQEGTLKQESSTEIPLQAESDQAAEEBAKQDEGEKEKEPTSPSPSSPVNSETTS 294
Qy 283 SPKKFTHGMAGRKKTSPFKKSKEDDLLETAKEKQEAQAEKVDSEKTEBPASEE---- 337
Db 295 TPKKFTHGMAGRKKTSPFKKSKEDDLLETAKEKQEAQAEKVDSEKTEBPASEE---- 354
Qy 338 -----QEPADTDQARLSADYKVELPLEDOYGDLEASSEKCAPLATEVPEKKEAHO-E 392
Db 355 QAPQEPBESAHBRLSLAEYKVELPLEDOYGDLEASSEKCAPLATEVPEKKEAHO-E 414
Qy 393 VVAEHNSTVEKTEBOGGGGAEGVVEGTGSLPPEKLAEPQVQEAPEELMK 452
Db 415 VVAEHNSTVEKTEBOGGGGAEGVVEGTGSLPPEKLAEPQVQEAPEELMK 466
Qy 453 REMCVSGGDHDTQLDLSPEEKLTPKHPEGIVSEVEMLSQGRIVQSPPLKLFSSGLK 512
Db 467 KETCVSGEDPTQADLSPDEKVLKSPPEGIVSEVEMLSQGRIVQSPPLKLFSSGLK 526
Qy 513 KLSGKKQKGRGGGDEPGEVYGHITESPESADQEGESSASSPEPEETTCLEKGL 572
Db 527 KLSGKKQKGRGGGDEPGEVYGHITESPESADQEGESSASSPEPEETTCLEKGL 585
Qy 573 APQGEAEETGTSDEGKKRREGITPMASFKKVTPKRRPSESDKEBELEKVASATLS 632
Db 586 VQDGEAEAGATSDGKKRREGITPMASFKKVTPKRRPSESDKEBELEKVASATLS 645

Qy 633 TDSTVSEMODEVKTVEBQKPEEKRRVDTSTVSEALICVSSKKRARKASSDDGCP 692
 Db 646 TESTASEMQEEMSGVSEPKPEEKRRVDTSTVSEALICVSSKKRARRSSDDEGPK 705
 Qy 693 TLGDSHRAEASKDKAGTDAVPASTQEQOQACSSSPPEAGSPSEGEVSTWESFKRL 752
 Db 706 AMGGDHQKADAGKDKGTGILAGSQEHDPGGSSPEAGSGFTGEVSTWESFKRL 765
 Qy 753 VTPRKKSGLKEKAED----SSVEQLSTEIEPSPRESWSIKKFIPORRKKRAGDKOEQ 808
 Db 766 VTPRKKSGLKEKSEDSIAGSGVSEHSPDTEPGKEESWSIKKFIPORRKKRPGDKOEQ 825
 Qy 809 ATVEDSGPVEINEDPNPAVPLSEYNAREKME---AQNTLPELLLCAGVVSSELS 865
 Db 826 APVEDAGPTGANEDSDVPAVPLSEYDAVERKMEAAQQAQGAEPQEKAAATEVSKELS 885
 Qy 866 KTLVHTSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEEYTEKDIIA-E 924
 Db 886 ESQVHMAAAVADGTRAATIIIEERSPSWISASVTEPLEOVEAEAAALLTEEVLEEVIAEE 945
 Qy 925 ETPVLTOTLPEGKAHDDMTSEVDFTSEAVTATETSEALTEEVTEASGAEETDMVSA 984
 Db 946 EPPTVTEPLPENREAGDTVVSEALTEPEAVTAETAETAGPLGSEEGTEASAAEETEMVSA 1005
 Qy 985 VSQLTSPDPTTEATPVQEVESGLVDTSEEROTQAILQAVDKVKEESQVPAT----- 1038
 Db 1006 VSQLTSPDPTTEATPVQEVESGLVDTSEEROTQAILQAVDKVKEESQVPAT----- 1065
 Qy 1039 -QTVORTGSKALEKVEEVEEDSEVLASEKXDMPKGVPQVQAGAEHLAQSGSETGOAPES 1097
 Db 1066 LQPVQRA-----FAERPEQAEASGLKTDVVLKVDQAQAKTEPFTQGVGQTTPE 1119
 Qy 1098 LE-VPEVTADVDH---VATCQV-----IKLQOL-MEQAVAPESSETLTDSETNGSTPLAD 1147
 Db 1120 FEKAPQVTESESSELVTTQAEITAGLVKSKQEMVMEQAI PPDVSVETPTDSETDGTSPVAD 1179
 Qy 1148 SDTAGTQODFTIISQSKATAAVRQSQVTEEBAAATAQKEPSPILPNNVPAQEBHEGEPG 1207
 Db 1180 FDAPGTTQKQDIVIHEENEVASGTSQSGTAEAAVPAQKERRPPA-PSSFVFEQETKEQSK 1238
 Qy 1208 -RDVLEPTQOELTAAAVPLAKTEVQSGEVDWLDGEKVK-----EEQEVFVHSG--PNS 1259
 Db 1239 MEDTLEHTDKVSVETVSIKSTEGTQ--EADQVADEKTKDVPFEGLEGSIDTIGITVSR 1296
 Qy 1260 QKAADVITYDSVMGAGQKESTEVQS-----LSLEEGEMETDVEKEKRETKPEQVSEE 1314
 Db 1297 EKVTEVALKGGTEEAECCKDDALELQSHAKSPSPVEREMVQVEREKTEAETHVNEE 1356
 Qy 1315 G-EQETAAPHEGYTKPVLTLDMPPSSRGKALGSLGS-PSLPDQDKAGCIEVQVQSLD 1372
 Db 1357 KLEHETAVTVSEEVSKQLLTQVNPFIIDGAKEVSSLEGSPPPCLGQEBQBAVCTKIQVQSS 1416
 Qy 1373 TTVTQTAEAQVKVI--ETVVISGESPCEVGAHLLPAEKSSATGGHWTQLQHAEDTVPGL 1430
 Db 1417 ASFLTAAAEKVKIGETANILETGETLEPAGAHVLVEEKSEKNEEDFAHPGSDAVPTG 1476
 Qy 1431 PESQAEPIIIVTAPESLTHPDQGEISASQORSESEEDKPDAGPDADGKESTAIERVL 1490
 Db 1477 PDCQAKSTFVIVSATTKGLSSDLEGEKTTSLKWKMSDEVQVQACQEV--KVSVAIEEDL 1534
 Qy 1491 KAEPEILESKSKNIVLNVITQAVDQPART-ETAPETHAVDSQTVQVPAKRLDSREPNRC 1549
 Db 1535 EPENGILELETKSKLVQNIITQAVDQVVRTEETATEMLTSELOTQAHVIRKADSDAQOE 1594
 Qy 1550 WTK-----MKDAKMKHPVQPPREDLQVLTLEA 1577
 Db 1595 TEKGGEDPQASQADETPTITSAKESESESTAVGQA 1627

RESULT 4
 AAO17365
 ID AAO17365 standard; protein; 1781 AA.
 XX

AAO17365;
 19-JUL-2002 (first entry)
 Human gravin.
 Human; endometriosis; DNA chip; fibronectin; p27; reticulocalbin;
 aldehyde dehydrogenase 6; gravin; phospholipase C epsilon; elastin;
 insulin-like growth factor binding protein-2; alpha-2 type IV collagen;
 transmembrane receptor PTK7; collagen type XVIII alpha 1;
 platelet derived growth factor receptor alpha; laminin M chain;
 subtilisin like protein PACB4; nidogen.
 Homo sapiens.
 EP1191107-A2.
 27-MAR-2002.
 21-AUG-2001; 2001EP-0250300.
 25-SEP-2000; 2000DE-1048633.
 (SCHD) SCHERING AG.
 Hess-stumpff H, Haendler B, Kraetzschmar J, Kreft B, Winterhager B;
 Regidor P, Scotti S;
 WPI; 2002-317413/36.
 In vitro diagnosis and monitoring of endometriosis, comprises
 detecting reduced expression of specific gene products, e.g. from the
 fibronectin gene -
 Claim 1; Page 18-19; 21pp; German.
 The present invention relates to a method for the in vitro diagnosis of
 endometriosis by determining the amount of gene product from at least one
 specific gene in a patient sample and comparing this with the amount of
 gene product in a control sample. A reduced level is indicative of
 endometriosis. The gene products may be fibronectin, p27, reticulocalbin,
 aldehyde dehydrogenase 6, gravin, phospholipase C epsilon, elastin,
 insulin-like growth factor binding protein-2, alpha-2 type IV collagen,
 transmembrane receptor PTK7, collagen type XVIII alpha 1,
 platelet derived growth factor receptor alpha, laminin M chain,
 subtilisin like protein PACB4 or nidogen. The method is useful for
 initial diagnosis of endometriosis, and also for monitoring progress and
 treatment of the disease. The present sequence is human gravin.
 Sequence 1781 AA;
 Query Match 53.2%; Score 4294; DB 23; Length 1781;
 Best Local Similarity 57.9%; Pred. No. 7.8e-221;
 Matches 959; Conservative 187; Mismatches 401; Indels 108; Gaps 35;
 Qy 1 MGAGSSTEQRSPSEQ-PAGSDTPSELVLSGHGPAAEAS-GAAGDPADA--DPATKLPQKNG 56
 Db 1 MGAGSSTEQRSPSEQPGSGSTPAEPSPGSGGSAEAPDTTADPAIAASDPATKLLQKNG 60
 Qy 57 QLSSVGVASQGDVHVQENQEQ-----EEVVDVEDVQRESEEDVREKDR 102
 Db 61 QLSTINGVABQDELSQLQGLDNGKQALNGQALNSQEEVEIVTEVGQDSEDSERDS 120
 Qy 103 VEEMAANSTAVEDITKDGQETSEIIEQIPASENNVMQVPAESQANDVGVKVFVFG 162
 Db 121 DKEMATKSAVVHDITDDQENRN-IEQIFSSSNLSELTQPTESQANDIGFKVVFVFG 179
 Qy 163 FKFTVKDKNEKSDTVQLLTVKKDEGEGBASVAGAGHQBPSPVETAVGESASKESELKQS 222
 Db 180 FKFTVKDKTEKPTVQLLTVKKDEGEGBA--AGAGDHQDPSL--GAGEAASKESEPKQS 234
 Qy 223 TEKQEGTLKQEQSSTEIPLOAESDQAAEEBAKDEGEKQKEPKTKSPESPSPVNSSETTS 282

```

Db 235 TEKEBETLKREQSAHEISPPABESGOAV-EECKEBSGEKQEKREPKSASPTSPVTSNGS 293
Qy 283 SFKKFFTHGNAMGRKKTSPFKSKSDEDDLETAERKKEQAEKVDEBEKEKTEBPASBE- 337
Db 294 TFKKFFPTQGMAGWRKKTSPFRPKPEDEVEASEKKQEPEKVDTEBDGKAFAVASEKLTA5E 353
Qy 338 ---QEPABEDTDQARLADYKVELPLEDQVGDLEASSEKCAPLATVPPREKMAHQ-E 392
Db 354 QAHPOEPABESAHBEPRLSAEYKVELPSEEOVSGSGPBEERAPAPLATEVPBEKLEVHOEE 413
Qy 393 VVAEVAHSTVKTETEEBOGGGAEBGGVVEGTESELPREKLAEPQEVQOEAAPAEELKMS 452
Db 414 VVAEVAHSTVETERTEEQ-----KTEVEETAGSVPAEBELVGMABEQAEPAEELVKL 465
Qy 453 REMCVSGGDHTQTLTDLSPBEKTLPRHPEGIVSEVEMLSQERIKVQSGPLKKLFSSGLK 512
Db 466 KETCVSGEDPTQAGADLSPEDEKVLSPREGVSEVEMLSQERIMKQVQGPLKLFSTGLK 525
Qy 513 KLSGKKQKRGKGGGDEPPEGYQHHTHSPESADQKQESSASSPEEBEETTCLEKPLE 572
Db 526 KLSGKKQKQGR-KGGDESGEHTQVPADSPSQEQKQESSASSPEEBEETTCLEKGLAE 584
Qy 573 APOGEAEEGTTSOGEKKREGTTPWASFKKMYTPKKRVRRPESPDEBELEKVKSGATLSS 632
Db 585 VQOGEAEEGATSDGEKKREGTTPWASFKKMYTPKKRVRRPESPDEBELEKVKSGATLSS 644
Qy 633 TDSITVSEMODEVKTGVEBOKPEEPKRRVDTSVSWEALICVSGSKKRRARAKASSDDDEGPR 692
Db 645 TESTISEMOEEMKGSVEEPKPEPKRKVDTSVSWEALICVSGSKKRRARRSSDDEGPRK 704
Qy 693 TLGGSDBRAEASAKQKXGTDVAVPASTOBDQAQSSSPPEPAGSSEBGVSTWMSFKRL 752
Db 705 AMGGHQADEAGKXKKEGTGDIAGSQEHDPQOQSSSPPEAGSPTEBGVSTWMSFKRL 764
Qy 753 VTPRKKSXSKLEKXAD-----SSVEQLSTELPSPRESVSTKKTIPGRKKRAGAKXQEO 808
Db 765 VTPRKKSXSKLEKXEDSIAGSGVHSTPDTBPKESVSTKKTIPGRKKRPRPKQEO 824
Qy 809 ATVEDSGVEINEDDPNPAVAVPLSEYNAVEREKME---AQNTLELPOLLGAVYVSEELS 865
Db 825 APVEDAGTGANEDSDVPAVAVPLSEYDAVEREKMEAOQAQGAQBPQKATVSKELS 884
Qy 866 KTLVHTSVAVIADTGRAVTSVEERSPSWISASVTPLEHTTAGAMPVPVEVTEKILIA-E 924
Db 885 ESQVHMAAAVADGTTRAATIIERSPSWISASVTPLEHTVENEALLTBEVLEVEREVIABE 944
Qy 925 ETPVLTOTLPBGKXADHDMVTSSEVDTSEAVATATETSEALTREEVTEASGAETTDWISA 984
Db 945 EPPVTEPLPENREARGTIVVSEALTPBAVTAETAGPLGSEEGTEASABEETTEMVSA 1004
Qy 985 VSQLTSDPTTEEATPVQEVESGLVDTSEERQTOALLQAVADKVEESQVPAE----- 1038
Db 1005 VSQLTSDPTTEEATPVQEVESGLVDTSEERQTOALLQAVADKVEESQVPAE----- 1064
Qy 1039 -QTVORTSKALEKVEEVBESVULASEKXQVMPKPVQEGAGHLAGSTGQATPES 1097
Db 1065 LQPVQRA-----EAERBEBOQAEGSLKKEITVVLKVAQOEAKTEPFTQKVGAGTTPES 1118
Qy 1098 LE-VEEVRADVDH---VATTCV-----IKLQOL-MEOAVAPESSELTIDSETNGSTPLAD 1147
Db 1119 FEKAOQVTESSSELVTTQOETLAGVKSQEMWEOALPPDSVETPTDSETDGSITPAVD 1178
Qy 1148 SPTADGTQODETIDSDSKATAAVRQSVTTEEATATQKEBESPTPNVPAOEBSGEFPG 1207
Db 1179 FDPARTTQKDEIVEIHEENEVAVSGTQSGGTEAFAPAQKERPPA-PSSPVPEBETKEQSK 1237
Qy 1208 -RDVLEPTQOELTAAAVPLAKTEVGOBGEVMDLGEKVK-----EBDEVFVHSG--PNS 1259
Db 1238 MEDLEHTDKAVSVETVSLSTEGNQ--EADQVADKTKDVPFEGLEGSIDTGITVSR 1295
Qy 1260 QKADVYTTDSEVMGVAQGEKESQVOS-----LSLEBEMTVDVKEKRETKPEOVSE 1314
Db 1296 EKVTEVALKGBETBAECKKDALQLQSHAKSPSPVEREMVVOVEREKTEAEPHVMEE 1355

```

```

Qy 1315 G-EOETAPAEHEGTYGKPVLTLMPSSEBKGALGSLGGS-PSLPDQKACIEVQVOSLD 1372
Db 1356 KLEHETIVTVSEEVSKQTLQTVNVPIDGAKEVSSLEGSPPCLGDEAVCTKIQVOSSE 1415
Qy 1373 TTVTQTAEAVEKVI--RTVVISSEBSPCEVCAGHLLPAEKSSATGHWTLQHAEDVPLG 1430
Db 1416 ASFTLTAAEBEKVLGTANILETGETLEPAGAHVLBEKSSSEKNEDEPAAHPGEDAVPTG 1475
Qy 1431 PESQAESPIIIVTPAPESTLHPDLQGEISASQERSEEDKPDAGPDADGKESVAILKVL 1490
Db 1476 PDCQASTPVIVSATFTKGLSSDLEGEKTSLSKWSDEVDVBOVACQEV--KVSVAIEDL- 1532
Qy 1491 KAEPB--ILHEESKNTIVANVLOTANDQFART-ETAPETHANDSQOVACRLDSEBPN 1547
Db 1533 --EPENGILELETSSKLQVNIITQAVDOFRTBETATEMLTSELQTOAHVIKADQSDAG 1590
Qy 1548 RCWTK-----MKDAKMGHPVQPREDLQVLTVLEA 1577
Db 1591 QETKEGEBPOASADETPTITSAKEBSESTAVGQA 1625

```

```

RESULT 5
AA53863
ID AA53863 standard; peptide; 1780 AA.
XX
AC AA53863;
XX
DT 13-JUL-1998 (first entry)
XX
DE Human gravin polypeptide.
KW Gravin; kinase anchoring protein; type II regulatory subunit; PKA; PKC;
KW cAMP-dependent protein kinase; protein kinase C; autoimmune disease;
KW Myasthenia gravis; nicotinic acetylcholine receptor.
XX
OS Homo sapiens.
XX
PN US5741890-A.
PD 21-Apr-1998.
XX
PF 19-DEC-1996; 96US-0769309.
PR 19-DEC-1996; 96US-0769309.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
PI Klauck TM, Nauert JB, Scott JD,
XX
DR WPI; 1998-260552/23.
XX
DR N-PSDB; AAV23545.
XX
PT New polypeptide fragments of protein kinase binding protein gravin -
PT are useful for the study of modulation of action between gravin and
PT protein kinase(s)
XX
PS Example 1; Column 19-32; 32pp; English.
XX
CC This sequence corresponds to the human gravin polypeptide, and represents
CC a polypeptide of the invention. The polypeptides are fragments capable of
CC binding to type II regulatory subunit of cAMP-dependent protein kinase
CC (PKA). Gravin is a kinase anchoring protein that binds to type II
CC regulatory subunits of PKA and protein kinase C (PKC). Gravin is also an
CC antigen of the autoimmune disease Myasthenia gravis (MG), where a patient
CC develops antibodies against their own nicotinic acetylcholine receptors.
CC The polypeptides are useful for providing analogues of gravin in the
CC study of the modulation (e.g. blocking, inhibiting and stimulating) of
CC interactions between gravin and kinase. The peptides are involved in the
CC modulation of gravin-kinase interactions.
XX
SQ Sequence 1780 AA;

```

Query Match	52.9%	Score 4274.5	DB 19	Length 1780
Best Local Similarity	57.8%	Pred. No. 8.5e-220		
Matches 957	Conservative 187	Mismatches 402	Indels 109	Gaps 36
Qy	1	MGAGSSTQRSPEQ--PAGSDTPSELVLSGHGPAARAS-GAAGDPADA--DPATKLPOKNG	56	
Db	1	MGAGSSTQRSPEQPPGSSTPAEPEPGGGSAAAPDPTADPAIAASDPATKLQKNG	60	
Qy	57	QLSSVGVVAEQGDVHVQBEHQEQ-----EEEVVDVGVQRESEDVREKDR	102	
Db	61	QLSTINGVAEQDELSQLQEGDLNGKALNGQALNASQREEEVIVTEVGQDSEDVSRDS	120	
Qy	103	VEEMAANSTAVEDITKGOQETSEIIEQIPASNNVEMVQPAESQANDVGPKVPKFGV	162	
Db	121	DKEMATKSAVVHDIITDDQEEENRN-IEQIPSSSENLLELTQPTESQANDIGPKVPKFGV	179	
Qy	163	FKFTVKDKNEKSDTVQLLTVKKDGEAGAAASVGAGDHOEPSVETAVGESASKESELQOS	222	
Db	180	FKFTVKDKTEKPDIVQLLTVKKDGEQA---AGADGHQDPSL--GAGEAASESEPKQS	234	
Qy	223	TEKQEGTLKQBSSTETIPLQASDOQAEEAAKDEGEKQEKETPKSPSPSPVNASSETTS	282	
Db	235	TEKPEETLKREQSHAEISPPAESQAV-EECKEEGEEKQEKPSKSAESPTSPVTSCTS	293	
Qy	283	SPKPFTHGWAGWRKKTSTFKSKS KEDDLTAEKRKQOEAKVDBEBKEKTEPASEE-----	337	
Db	294	TFKPFTHGWAGWRKKTSTFRPKPEKDEVEASEKKQEPEKVDTEBDGKAEVASEKLTASE	353	
Qy	338	---QEPAREDQQARLSADYKEVELPLEDQCDLSEASSEKCAPLATVDFDEKMAHQ-E	392	
Db	354	QAHPQEPAESAHPEKLSAEBYKEVELPSEBQVSGSGPSEKPAPLATEVDFDKIETHQBE	413	
Qy	393	VVAEVHVSTVEKTEEBEQGGGAEGGVVVGTEGSLPPEKLAEPQEVPOEAPPAEELMKS	452	
Db	414	VVAEVHVSTVEERTEQ-----KTEVEETAGSVPAEELVGMDAEPOEAPPAKELVKL	465	
Qy	453	REMCVSGGDHTQLTDLSPBEKTLPKHPGPIVSEVEMLSQERIKVQGSPLKKLPSSGLK	512	
Db	466	KETCVSGDPTQAGDLSDPKLVSKPPGCVSEVEMLSQERMKVQGSPLKKLFTSTGLK	525	
Qy	513	KLSGKKQKRGKGGDEBPGEYQHHTSPESADQKSGSSASSPEEPEETCLEKGPLE	572	
Db	526	KLSGKKQKGRK-GGGDEBSGHQTPADSPDSPEQKSGSSASSPEEPEETCLEKGLAE	584	
Qy	573	APQGEABEGTTSQGEKKREGITPWASPKMVTPKKRVRRPSESDEKELEKVKKSATLSS	632	
Db	585	VQQGEAEAGATSDGEKKREGVTPWASPKMVTPKKRVRRPSESDEKDELKVKKSATLSS	644	
Qy	633	TDSTVSEMQDEVKTVGEQKPEEPKRRVDTSVSWEALICVGSKKRKAASSSDDGGPR	692	
Db	645	TESTASEMOEBMKGSEBPKEPKEPKKVDTSVWEALICVGSKKRARRSSSDEGGPK	704	
Qy	693	TLGDSHRAEASDKKEAGTDAVPASTQEQDAQGSSSPGAPGSPGEGVSTWESFKRL	752	
Db	705	AMGSDHQKADAGKDKETGTGILAGSQEHDPGQSSSPGAPGSPGEGVSTWESFKRL	764	
Qy	753	VTPRKKS KLEEKAEAD-----SSVEQLSTETSPSEESVWSIKKPIGRKKRKAQKQEQ	808	
Db	765	VTPRKKS KLEEKSEDSIAGSGVBSHTDTPTEPGKESVWSIKKPIGRKKRPPGKQEQ	824	
Qy	809	ATVEDSGPVEINDDPNYPVAVPLSEYNVAREKME---AQNTLPLQLLGAIVVSEELS	865	
Db	825	APVEDAGPTGANEDSDVPVAVPLSEYDAVEREKMEAQAQKGAQEPQKAAATEVSKELS	884	
Qy	866	KTLVHTSVAVIDGTRAVTSVEERSPSWISASVTBPLBHTAGEAMPVVEVTEKDIIA-E	924	
Db	885	ESQVHMMAAAVADGTAAATIIIEERSPSWISASVTBPLBQVEAAEALLTEEVLEREVIABE	944	
Qy	925	ETPVLTOTLPGKDAHDDMWTSVDFTSEAVTATSTSEALRTEEVTEASGAEETDMVSA	984	
Db	945	EPPIVTEPLPENARGTOTVSEAEITPEAVTAAETAGTGLGSEEGTASAABETTEMVSA	1004	
Qy	985	VSQLTDSPTTTEATPPQVEBSGVLDTEEBERQTOAILQAVADVKKESQVPAAT-----	1038	

PR 19-DEC-1996; 96US-0769309.
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 XX KLauck TM, Scott JD, Nauert JB;
 XX WPI; 2000-523763/47.
 DR N-PSDB; AAA74903.
 XX Novel polynucleotides useful for detecting gravin in patients suffering
 PT from Myasthenia gravis encodes CAMP-dependent protein kinase-binding
 PT polypeptide and protein kinase C-binding polypeptide of gravin -
 XX
 PS Claim 1; Column 35-45; 34pp; English.
 XX
 CC The present sequence is the protein sequence of human gravin. Gravins is
 CC an A-kinase anchoring protein (AKAP) which is involved in the
 CC localisation of CAMP dependent protein kinase A (PKA) via interactions
 CC between the RII binding region and the PKA regulatory subunit RII. Gravins
 CC is also an antigen found in myasthenia gravis sufferers, and it is
 CC thought that antibodies to it may be useful in modulating the binding of
 CC PKA, and thus aid in the treatment of the disease. The gravin coding
 CC sequence was isolated by first screening a human umbilical vein
 CC endothelial cell cDNA library with serum from a myasthenia gravis
 CC patient, and then searching a human heart cDNA library for sequences
 CC resembling the isolated sequence. This was done because the first
 CC sequence obtained was shown to be shorter than the full length cDNA.
 XX
 SQ Sequence 1780 AA;
 Query Match 52.9%; Score 4274.5; DB 21; Length 1780;
 Best Local Similarity 57.8%; Pred. No. 8.5e-220;
 Matches 957; Conservative 187; Mismatches 402; Indels 109; Gaps 36;

QY 573 APDGEAEGETTSDGKKRRCITPMASFKKMTPKKRVRRPSSDKEBELEKYSATLSS 632
 DB 585 VQDGEAEGETTSDGKKRRCITPMASFKKMTPKKRVRRPSSDKEBELEKYSATLSS 644
 QY 633 TDSTVSEMODEVKTGVEEQKPEEPKRRVDTSVSWEALICVSSSKKARKASSDDEGGPR 692
 DB 645 TESTASMODEEMKGVSEPRPEEPKRRVDTSVSWEALICVSSSKKARKARRSSDDEGGPK 704
 QY 693 TLGDSHRAEASDKKAGTDAVPASTQEDDQAGSSPEPPASPSSEGEVSTWESKRL 752
 DB 705 AMGDHQAEDAGKDKETGDTGILAGQEHDPGQSSSPQAGSPTEGEGVSTWESKRL 764
 QY 753 VTPRKKSKSLKEEKAED---SSVEQLSTELPSPRSWSWISIKKIFGRKKADAGQEQ 808
 DB 765 VTPRKKSKSLKEEKSSEDSINGSGVEHSTPTEPKEKSWWSIKKIFGRKKRPDQEQ 824
 QY 809 ATVEDSGPVEINEDDPVAVPVPVLESEYNAVERKME---AQNTLELPOLLGAVVSELS 865
 DB 825 APVEDAGPTGANDSDVPVAVPVPVLESDAVREMEQAQKAGQEQKAAATEVSKELS 884
 QY 866 KTLVHTVSAVIDGTAVTSVEERSPSWISASVTEPLEHTAGSAMPVVEYTEKDITL-E 924
 DB 885 ESQVHMAAAVAADGTAAITIEERSPSWISASVTEPLEQVEAEALTLTEVLEREVIAEE 944
 QY 925 ETPVLQTLPEGKADHDMTSEVDFTESEAVTATESEALRTEVTEASGAEEETDQVSA 984
 DB 945 EPTVTEPLEPNNRRARGDTVSEBALTEPVAITAEITGPIQSEEGTASAAEETTEMVSA 1004
 QY 985 VSQLTDSPTTEEATPVQVEESGVLDTEEEERQTAIIOAVADVKKESQVPAT----- 1038
 DB 1005 VSQLTDSPTTEEATPVQVEESGVLDTEEEERQTAIIOAVADVKKESQVPAT----- 1064
 QY 1039 -QTVQRTSKALEKVEEVEEBSVLASBKEDVMPKGPVQEAAGNHLAQSGETQAPSES 1097
 DB 1065 LQPVQRA-----BAEREEQAEASGLKETDVVLKXDAQAKTEPTQGVVQQTTPES 1118
 QY 1098 LE-VPEVADVDH---VATQGV-----IKLQOL-MEOAVASESETLIDSETNSTPLAD 1147
 DB 1119 FEKAPQVTESESELVTTCAETLAGVQSEWMEQALIPDSVETPTDSTSTDSPTPAD 1178
 QY 1148 SPTADGTQODETIDSDSKATAVAQSQVTEEEBAATQKEEPSTLPNNVPAQEBHGEPG 1207
 DB 1179 FPAEGTQKDELVEIHEENENHLV-PVRGTEAEAVPAQKEPRA-PSSFVQOETKEQSK 1236
 QY 1208 -RDVLEPTQOELTAAAVLVLAKTVEQGEVDMLDEKVK-----EBOEVFVHSG--PNS 1259
 DB 1237 MEDTLEHTDKEVSVETVAILSTKEGTQ--EADQYADEKTKQVPPFEGLEGSIDTGITVSR 1294
 QY 1260 QKADAVTDTSEVMGAVQGEKESTEVQS-----LSLEEGMETVDEKERTKPEQVSE 1314
 DB 1295 EKVTVALKSGTEBAEKQKDALIELQSHAKSPSPVRENNVQVERKTEAEHPHVAEE 1354
 QY 1315 G-EQETAAPENHGTVGKGVLTLDMPSSERKALGSLGS-PSLPDQDAGCIEVQVQLD 1372
 DB 1355 KLEHTATVTVSEVSKQLQTVNVPIDGAKESVLSBSPPCGQGEAVCTKIQVQSSB 1414
 QY 1373 TTVTQTAABEAKVTL-ETVVISSEGESPECVGAHLLPAEKSSATCGHMTLQHAETVPLG 1430
 DB 1415 ASFTLTAAABEAKVLTGFTANILETGETLEPAGAHVLVEKSESEKNEEDPAHGPDEAVPTG 1474
 QY 1431 PESQAESIPITTPAPBESTLHDLQELISASQRESESEEDKPDADPDGKSTAIKYL 1490
 DB 1475 PDCQAKSTFVIVASATTKGLSSDLGEEKTSLKMSDSEVDQVACQEV-KVSAVIEDL- 1531
 QY 1491 KAEPE--ILELESKNKIYLVNITQAVDQFART-ETABETHAVDSQOTQVAPACRLDSREBN 1547
 DB 1532 --EPENGILEETKSKLVQNIQITAVDQFVTEETATHEMLTSELQTAHVITKADSDQAG 1589
 QY 1548 RCMWK-----MKDAKQKHPVQPRFDLOVLYVLEA 1577
 DB 1590 QETEKKEGEPEQASQADEPTITSKESESESTAVGOA 1624


```

QY 1148 SDPADGTOQDETTIDSDQSKATAAVRSQYTHEEAATPAKKEPSTLPPNNVPAQOEHEBPG 1207
Db 1188 FDABGTTQKQKEIYIHEHENEVASGTQSGGEAAVPAQKEPRA-PSSFVQEBTEKQSK 1246
QY 1208 -RDVLEPFOOEILTAAPVAVIAKTEVQOEGBEVMLDGEKYK-----EGBEVEVHSG--PNS 1259
Db 1247 MEDLEHTIDKEVSEVYISLTKNEGQ--EADQYADEKQKDPFPELEGSIDPTIVSR 1304
QY 1260 QKADVTYDSEWGVAGCCQKESTEVQS-----LSIEBEMETDYERKRETK--PEQVSE 1313
Db 1305 EKVLEVVAIKEGTEAEACKDDALEQSHAKSPSPSEVHEMVQYEREYVQXOSPETHNE 1364
QY 1314 EG-BOETNAPEHECTGYKPAULTIDMPSSEKRALGSLGS--PSLPDQKACIIVQYQSL 1371
Db 1365 EKLHEHTNAVYSEVQSQOLLQTNVPIIDGAKEVSLBESPPCLGQGEAATCMIQVQSS 1424
QY 1372 DTVTQTAEAVERKI--ETVVISETGSEPCVGAHLLPAKKSATGTGHTLQHAEDTVPL 1429
Db 1425 EASFLLTAAAEKEKVLGETANILETGETLEPAGAHVILEKSEKNEDPAHPGEDVAPT 1484
QY 1430 GPESGABEIPITIVPAEESTLHPDLOGELISASQREHSEEDKPDGADGADGESATIEKY 1489
Db 1485 GPDCQASTPIVATTKKGLSSDLBEKTKTSLKMSXSDVDVQVACQEV--KVSVALBED 1542
QY 1490 LKAEPEILELESKNKTLVNVLOTAVDQFART-ETAPETHAVDSQTOVPACRLDSREBNR 1548
Db 1543 LEPENGILEETKSKLVQNIIGTAVDQFRTETETEMLTSELQTAHVIAKDSQDAGQ 1602
QY 1549 CWTK----MDAKMKHPVDPQPREDDLOLVTLVLEA 1577
Db 1603 ETEKEGEDPQASADETPITSABEESSTAVGQA 1636

```

AAW53862	standard; peptide; 292 AA.
AAW53862	
AAW53862	
13-JUL-1998	(first entry)
Gravin polypeptide fragment, residues 265-556.	
Gravin, kinase anchoring protein; type II regulatory subunit; PKA; PKC; cAMP-dependent protein kinase; protein kinase C; autoimmune disease; Myasthenia gravis; nicotinic acetylcholine receptor.	
Homo sapiens.	
US5741890-A.	
21-APR-1998.	
19-DEC-1996;	96US-0769309.
19-DEC-1996;	96US-0769309.
(UYOR-) UNIV OREGON HEALTH SCI.	
Klauck TM, Nauert JB, Scott JD;	
WPI; 1998-260552/23.	
New polypeptide fragments of protein kinase binding protein gravin - are useful for the study of modulation of action between gravin and protein kinase(s)	
Claim 3; Column 17-20; 32pp; English.	
This sequence corresponds to residues 265-556 of gravin, and represents a polypeptide of the invention. The polypeptides are fragments capable of binding to type II regulatory subunit of cAMP-dependent protein kinase	

CC (PKA). Gravin is a kinase anchoring protein that binds to type II
CC regulatory subunits of PKA and protein kinase C (PKC). Gravin is also an
CC antigen of the autoimmune disease Myasthenia gravis (MG), where a patient
CC develops antibodies against their own nicotinic acetylcholine receptors.
CC The polypeptides are useful for providing analogues of gravin in the
CC study of the modulation (e.g. blocking, inhibiting and stimulating) of
CC interactions between gravin and kinase. The peptides are involved in the
CC modulation of gravin-kinase interactions.
XX
XX
SQ Sequence 292 AA;

Query Match 11.8%; Score 952.5; DB 19; Length 292;
Best Local Similarity 66.1%; Pred. No. 2.7e+43;
Matches 199; Conservative 30; Mismatches 53; Indels 19; Gaps 4

```

QY 254 K0EGEK0EK0EPTYSPBSPSSPVNSETTSSPPKPFPTTGMAGMKKTSPKSKEDDLETAB 313
Db 1 KEEGEEK0EK0EPTSAESPTSPVNSESTGSPFKKFFPTGGMAGMKKTSPKPKDEVEAS 60
QY 314 KKK0EAKVDBEEKETEPSE-----QEPBDDQARLADYKXELPLEDOY 364
Db 61 KKK0E0EPKVT0EEDGAEVASEKLTLS00ANP0EP0EASHEPRLTSAEYKVELPSE0Y 120
QY 366 GDLEASSEK0APLTATEVDFDKM0AN0-EVYAEVHYATVTEKTE00GGGAEAGVVEG 423
Db 121 SGG0SPSE0EKPAPLATEVDFDKLTVH0EBVVAEYHVSIVERTBEO-----KTEVEE 172
QY 424 TGESLPPEKLAPOEVP0EAPAEELMKSRM0CVSGDHTLDTLSPEEKTLPKH0EGIV 483
Db 173 TAGSVPAEELVGM0AP0EAPAEELVYKLT0CVSGSDPT0GADLSDEKVL0SKP0E0V 222
QY 484 SEVENMSS0ERIKV0GSDPLKTLFSSSGLKTL0GKK0KKG0GGGDEE0PGEY0HINTESP 543
Db 223 SEVENMSS0ERIKV0GSDPLKTLFSTGHLKTL0GKK0GKR-GGGDEESGHT0VPADSPD 291
QY 544 S 544
Db 292 S 292

```

```

XX RESULT 9
XX AAB15379
XX ID AAB15379 standard; protein; 292 AA.
XX AC AAB15379;
XX DT 26-JAN-2001 (first entry)
XX DE Recombinant human gravin polypeptide fragment #1.
XX KW Human; gravin; PKA RII binding site; myasthenia gravis;
XX KM kinase anchoring protein; CAMP dependent protein kinase.
XX OS Homo sapiens.
XX PN US6090929-A.
XX PD 18-JUL-2000.
XX PF 19-DEC-1997; 97US-0994570.
XX PR 19-DEC-1996; 96US-0769309.
XX PT (UYOR-) UNIV OREGON HEALTH SCI.
XX PA Klaufack TM, Scott JD, Nauert JB;
XX PI WPI; 2000-523763/47.
XX DR
XX PT Novel polynucleotides useful for detecting gravin in patients suffering
XX from Myasthenia gravis encodes CAMP-dependent protein kinase-binding
XX PT polypeptide and protein kinase C-binding polypeptide of gravin -

```

PS Example 4; Column 21-24; 34pp; English.

XX The present sequence is a recombinant version of the human gravin

CC protein. Gravin is an A-kinase anchoring protein (AKAP) which is involved

CC in the localisation of cAMP dependent protein kinase A (PKA) via

CC interactions between the RII binding region and the PKA regulatory

CC subunit RII. Gravin is also an antigen found in myasthenia gravis

CC sufferers, and it is thought that antibodies to it may be useful in

CC modulating the binding of PKA, and thus aid in the treatment of the

CC disease. This sequence was used in an assay to determine the ability of

CC gravin to bind protein kinase C.

XX Sequence 292 AA;

Query Match 11.8%; Score 952.5; DB 21; Length 292;

Best Local Similarity 66.1%; Pred. No. 2.7e-43;

Matches 199; Conservative 30; Mismatches 53; Indels 19; Gaps 4;

QY 254 KDGEEKQKEPTKSPSPVNSSETTSFKKFFTHGAGWRKTSFKKSKEDDLTAETAE 313

DB 1 KEGEEKQKEPKSAESPTSPVTSSTGTFKFFTHGAGWRKTSFKRKPXSDEVEASE 60

QY 314 KRKEQAEKVDDEEKEKTEPASEE-----QEPADTDQARLSADYEKVELPLEDOV 364

DB 61 KRKEQEPKVDTEEDGRAEVAEKLTASEQAHQPEAPSAEHPRLSAEYKVELPSEQV 120

QY 365 GDLEASSEEKCAPLATVDFEKEMAHQ- EVVAHVHSTVTEKEEQGGGEAGGVVVEG 423

DB 121 SGSGPSEKPAPLATEVDFEKEVHQEVEVVAHVHSTVTEEQ-----KTEVEE 172

QY 424 TGESLPKEKLAEPQEPQAEPAELMKREMCVSGGDHQLTDLSPEEKTLKHPGIV 493

DB 173 TAGSVPAELVGMDAEPQAEPAELKVLKETCVSGEDPTQGDALSPDEKVLKPPPEGV 232

QY 484 SEVEMLSQERIKVQGSPLKLFSSGLKLGKQKKGKGGGDEBPGEVQHIHTESPE 543

DB 233 SEVEMLSQERIKVQGSPLKLFSTGLKLGKQKKGK- GGDDESGEHTQVPADSPD 291

QY 544 S 544

DB 292 S 292

RESULT 10

AAW53875

ID AAW53875 standard; peptide; 651 AA.

AC AAW53875;

XX 13-JUL-1998 (first entry)

DE Gravin polypeptide fragment, residues 1130-1780.

XX Gravin; kinase anchoring protein; type II regulatory subunit; PKA; PKC;

KW cAMP-dependent protein kinase; protein kinase C; autoimmune disease;

KW Myasthenia gravis; nicotinic acetylcholine receptor.

OS Homo sapiens.

XX US5741890-A.

XX 21-APR-1998.

XX 19-DEC-1996; 96US-0769309.

XX 19-DEC-1996; 96US-0769309.

XX (UYOR-) UNIV OREGON HEALTH SCI.

XX Klauck TM, Nauert JB, Scott JD;

XX WPI; 1998-260552/23.

XX

PT New polypeptide fragments of protein kinase binding protein gravin -

PT are useful for the study of modulation of action between gravin and

XX protein kinase(s)

XX Example 4; Column 49-54; 32pp; English.

XX This sequence corresponds to residues 1130-1780 of gravin, and represents

CC a polypeptide of the invention. The polypeptides are fragments capable of

CC binding to type II regulatory subunit of cAMP-dependent protein kinase

CC (PKA). Gravin is a kinase anchoring protein that binds to type II

CC regulatory subunits of PKA and protein kinase C (PKC). Gravin is also an

CC antigen of the autoimmune disease Myasthenia gravis (MG), where a patient

CC develops antibodies against their own nicotinic acetylcholine receptors.

CC The polypeptides are useful for providing analogues of gravin in the

CC study of the modulation (e.g. blocking, inhibiting and stimulating) of

CC interactions between gravin and kinase. The peptides are involved in the

CC modulation of gravin-kinase interactions.

XX Sequence 651 AA;

Query Match 8.1%; Score 656; DB 19; Length 651;

Best Local Similarity 38.3%; Pred. No. 6.1e-27;

Matches 191; Conservative 80; Mismatches 188; Indels 40; Gaps 17;

QY 1110 VATCOV-----IKLQQL-MEOAVAPESSETLTDSETNGSTPLADSDTADGTQODEIDSQ 1163

DB 6 VTTQCAETLAGVQSBVMVMEQAIIPDPSVETPTDSETDGTGTPVADFAPGTQKDEIVEIH 65

QY 1164 DSKATAAARQSOVSTEEAATAOKEEPTLNNVPAQEEHGEPEG-RDVLPTQOELTAAA 1222

DB 66 ENEVELV-PVRGTEAEVPAQKERPPA-PSSFVFOEETKEQSKMEDTLEHTDKESVET 123

QY 1223 VPVLAKEVQGEVDMLDGKVKV-----EEQEVFVHSG--PNSQKAADVTDSEYMGVA 1275

DB 124 VSILSKTEGTQ--EADQYADEKTKDVPFFEGLEGSIDTGTIVSRKRVTEVALKGEGTEEA 181

QY 1276 GCOEKESTEVQS-----LSLEGEMETDVEKRETKPQVSEEG-EQETAAPHEHGTYG 1329

DB 182 ECKKDDALELQSHAKSPSPFVEREMVYQVEREKTEAEPTHVNEEKLHETAVTVSEVSK 241

QY 1330 KPVLTLDMPSSSRGKALGSLGGS-PSLPDQKAGCTEVQVSLDITVTOTAAEAVEKVI-- 1386

DB 242 QLLQTVNVPIIDGAKESVSLGSPPPCLGQEEAVCTKIQQVSEASEFTLITAAAEKVLG 301

QY 1387 ETVVISSETGESPECVGAHLLPAEKSSATGCHWTLOHAEDTVPGLGPSQAESIPPIVTPAP 1446

DB 302 ETANILETGETLEPAGAHVLEEKSEKNEDEFAAHGAGEDAVPTGPDCAKSTPVIYSATT 361

QY 1447 ESTLHPDLOGEISASQERSEEDKPDAGPDADGKESSTAIEKVLKAEPE--ILELESKN 1504

DB 362 KKGSLSDLEGEKTTSLKWSKDEVDQVACQEV--KVSVAIEDL---EPENGILELETKSS 416

QY 1505 KIVLVNIQTAVDQFART-ETAPETHAYDSQTOVPACRLDSRENRCTWK----MKDAKM 1558

DB 417 KLVQNIQTAVDQFVTEETATMLTSELQTAHVIKADSDQAGQETEKTEKEEPPQASAOQ 476

QY 1559 KHEVPQPFREDLQVLTVLEA 1577

DB 477 ETPITSAKESESESTAVGQA 495

RESULT 11

AAAB18410

ID AAAB18410 standard; protein; 651 AA.

XX AAAB18410;

XX 26-JAN-2001 (first entry)

XX Human gravin PKA regulatory subunit RII binding site assay peptide #3.

XX Human; gravin; PKA RII binding site; myasthenia gravis;

KW kinase anchoring protein; cAMP dependent protein kinase.

PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX Klauck TM, Nauert JB, Scott JD;
 XX WPI, 1998-260552/23.
 DR
 XX
 PT New polypeptide fragments of protein kinase binding protein gravin -
 PT are useful for the study of modulation of action between gravin and
 PT protein kinase(s)
 XX
 PS Example 2; Column 43-46; 32pp; English.
 XX
 CC This sequence corresponds to residues 1130-1582 of gravin, and represents
 CC a polypeptide of the invention. The polypeptides are fragments capable of
 CC binding to type II regulatory subunit of cAMP-dependent protein kinase
 CC (PKA). Gravins is a kinase anchoring protein that binds to type II
 CC regulatory subunits of PKA and protein kinase C (PKC). Gravins is also an
 CC antigen of the autoimmune disease Myasthenia gravis (MG), where a patient
 CC develops antibodies against their own nicotinic acetylcholine receptors.
 CC The polypeptides are useful for providing analogues of gravin in the
 CC study of the modulation (e.g. blocking, inhibiting and stimulating) of
 CC interactions between gravin and kinase. The peptides are involved in the
 CC modulation of gravin-kinase interactions.
 CC
 XX Sequence 453 AA;
 SQ
 Query Match 7.9%; Score 639.5; DB 19; Length 453;
 Best Local Similarity 40.5%; Pred. No. 2.8e-26;
 Matches 183; Conservative 69; Mismatches 165; Indels 35; Gaps 16;
 QY 1110 VATCOV-----IKLOOL-MEOAVAPESSETLTDSENGSTPLADSDPTADGQOEDTDSQ 1163
 DB 6 VTTQCAETLAGVKSGEMVMEQAIIPDSVETPTDSETDSTVADFPDAGCTQKOEIYVH 65
 QY 1164 DSKATAVROQVTEEBEATAQKEBPTLPNNVPAQEEHGEPPG-RDYLEPTQOELTAA 1222
 DB 66 ENEVHLV-PVRGTEAEAVPAQKERPPA-PSSFVFOEETKEQSKMEDTLEHTDKEVSVE 123
 QY 1223 VPVLAKTEVGOGEVDWLDGKVK-----EEOEYFVHSG--PNSQKADVYDSEVMGYA 1275
 DB 124 VSILSKTEGTQ--EADQYADEKTKDVPFEGLEGSGIDGTITVSRKVTVALKGEETBEA 181
 QY 1276 GGOEKESTVOS-----LSLEGEMETDVEKEKETKEQVSEEG-EOETAPEHEGTYG 1329
 DB 182 ECKKDALELOSHAKSPSPVEREMVQVREKTEAEFTHNEEKLEHTAVTSEBVS 241
 QY 1330 KPVLTLDMPSSERKALGSLGGS-PSLPDQKAGCIEVOVSLDTTQTQTAABAEKVI-- 1386
 DB 242 QLIQIVNPIIDGAKESVLSGSPPCLOGEBAVCTKIQVOSSEASFTLTAABEEKVIG 301
 QY 1387 ETVVISSETGESPECVGAHLIPAESSATGCHWTLOHAEDTVPLGEGSAESIPITVPAP 1446
 DB 302 ETANILETGETLEPAGAHVLVEKSSSEKNEDEFAHPGEDAVPTGPDCAKSTPVIVSAT 361
 QY 1447 ESTLHPDLOGEISASQRESEEDKPDADGKSTAEIKVLAEBE--ILLEESKSN 1504
 DB 362 KKGSLSDLEGETTSLKMXSDVEDEQVACOEV--KVSVAIEDL--EPENGILELETSS 416
 QY 1505 KIVLVNIQTAVDOFART-ETAPETHAYDSQTQ 1535
 DB 417 KLVONIIQTAVDOFVTRTEETATEMLTSELQTQ 448

RESULT 14
 AAB18407
 ID AAB18407 standard; protein; 453 AA.
 XX
 AC AAB18407;
 XX
 DT 26-JAN-2001 (first entry)
 XX
 DE Human gravin PKA regulatory subunit RII binding site assay peptide #1.
 XX

KM Human; gravin; PKA RII binding site; myasthenia gravis;
 KM kinase anchoring protein; cAMP dependent protein kinase.
 XX
 XX Homo sapiens.
 OS
 XX US6090929-A.
 PN
 XX 18-JUL-2000.
 PD
 XX 19-DEC-1997; 97US-0994570.
 PF
 XX 19-DEC-1996; 96US-0769309.
 PR
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX Klauck TM, Scott JD, Nauert JB,
 PI WPI, 2000-523763/47.
 DR
 XX
 PT Novel polynucleotides useful for detecting gravin in patients suffering
 PT from Myasthenia gravis encodes cAMP-dependent protein kinase-binding
 PT polypeptide and protein kinase C-binding polypeptide of gravin -
 CC
 CC
 XX Example 2; Column 47-50; 34pp; English.
 XX
 CC The present sequence is part of the human gravin protein. Gravins is an
 CC A-kinase anchoring protein (AKAP) which is involved in the localisation
 CC of cAMP dependent protein kinase A (PKA) via interactions between the
 CC RII binding region and the PKA regulatory subunit RII. Gravins is also an
 CC antigen found in myasthenia gravis sufferers, and it is thought that
 CC antibodies to it may be useful in modulating the binding of PKA, and thus
 CC aid in the treatment of the disease. This sequence was used in an assay
 CC to identify the precise position of the RII binding site within the
 CC gravin protein.
 CC
 XX Sequence 453 AA;
 SQ
 Query Match 7.9%; Score 639.5; DB 21; Length 453;
 Best Local Similarity 40.5%; Pred. No. 2.8e-26;
 Matches 183; Conservative 69; Mismatches 165; Indels 35; Gaps 16;
 QY 1110 VATCOV-----IKLOOL-MEOAVAPESSETLTDSENGSTPLADSDPTADGQOEDTDSQ 1163
 DB 6 VTTQCAETLAGVKSGEMVMEQAIIPDSVETPTDSETDSTVADFPDAGCTQKOEIYVH 65
 QY 1164 DSKATAVROQVTEEBEATAQKEBPTLPNNVPAQEEHGEPPG-RDYLEPTQOELTAA 1222
 DB 66 ENEVHLV-PVRGTEAEAVPAQKERPPA-PSSFVFOEETKEQSKMEDTLEHTDKEVSVE 123
 QY 1223 VPVLAKTEVGOGEVDWLDGKVK-----EEOEYFVHSG--PNSQKADVYDSEVMGYA 1275
 DB 124 VSILSKTEGTQ--EADQYADEKTKDVPFEGLEGSGIDGTITVSRKVTVALKGEETBEA 181
 QY 1276 GGOEKESTVOS-----LSLEGEMETDVEKEKETKEQVSEEG-EOETAPEHEGTYG 1329
 DB 182 ECKKDALELOSHAKSPSPVEREMVQVREKTEAEFTHNEEKLEHTAVTSEBVS 241
 QY 1330 KPVLTLDMPSSERKALGSLGGS-PSLPDQKAGCIEVOVSLDTTQTQTAABAEKVI-- 1386
 DB 242 QLIQIVNPIIDGAKESVLSGSPPCLOGEBAVCTKIQVOSSEASFTLTAABEEKVIG 301
 QY 1387 ETVVISSETGESPECVGAHLIPAESSATGCHWTLOHAEDTVPLGEGSAESIPITVPAP 1446
 DB 302 ETANILETGETLEPAGAHVLVEKSSSEKNEDEFAHPGEDAVPTGPDCAKSTPVIVSAT 361
 QY 1447 ESTLHPDLOGEISASQRESEEDKPDADGKSTAEIKVLAEBE--ILLEESKSN 1504
 DB 362 KKGSLSDLEGETTSLKMXSDVEDEQVACOEV--KVSVAIEDL--EPENGILELETSS 416
 QY 1505 KIVLVNIQTAVDOFART-ETAPETHAYDSQTQ 1535
 DB 417 KLVONIIQTAVDOFVTRTEETATEMLTSELQTQ 448

RESULT 15
 ABB66811
 ID ABB66811 standard; Protein; 6815 AA.
 XX AC ABB66811;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 27225.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX DR WPI; 2001-656860/75.
 XX DR N-PSDB; ABL10914.
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
 XX PT interactions -
 XX PS Disclosure; SEQ ID NO 27225; 21pp + Sequence Listing; English.
 XX CC The invention relates to an isolated nucleic acid detection reagent
 XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
 XX CC useful in developmental biology and in elucidating cell signalling and
 XX CC cell-cell interactions in higher eukaryotes for the development of
 XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
 XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 XX CC sequences (ABL01840-ABL16175) and the encoded proteins
 XX CC (ABB57737-ABB72072).
 XX CC The sequence data for this patent did not form part of the printed
 XX CC specification, but was obtained in electronic format directly from WIPO
 XX CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 6815 AA;

Query Match 6.9%; Score 557.5; DB 22; Length 6815;
 Best Local Similarity 20.5%; Pred. No. 3.1e-20;
 Matches 393; Conservative 304; Mismatches 716; Indels 503; Gaps 80;

QY 7 TEQSPQAG-----SDTP-SELVLSGHGPAAEASGAAGPADADPATKLPQK 54
 DB TTKMLRRPAGEIEIIEVVRDQPEAEITVEYEP-----EPVQDEKPEPKK 3039
 QY 55 NGQLSSVNGVAEQGDVH--VQ-----BENQEQEEVVDE-----DVQRE 93
 DB KTR-----KVKDDIHDIQKLEIETPKTELEKYEKIEPIVKDKPLDSPDLDES 3093
 QY 94 SEDVREKDR-----VREMANSTAVEDITKQGEETSEIEIQIPASENNVEE 140
 DB PKEVQKDKKSRSTKVNEETPVQEQAKVNVVEEAPEQEPFVQILEVKPV-EVDVKE 3152
 QY 141 MVQPAESQANDVGKPKFVYG-----FKFTVKDKNEKSDTVQLLTVKKDEGEAEAS 194
 DB VITEDGKPVQKTKTVLKKLGIPQEOTFKITM--ISEDNDNV---TVIVDEPEETASP 3207
 QY 195 VGAGDHQPSVETAVGESAS-----KESEL-----KQSTEKQEGTLKQEQSS 236

DB 3208 QSIEEHPEQSKELAPKPKTKTVRKVKDDLSYVKKLIEEIEPKVDLEKYE---KVEMPE 3264
 QY 237 TEIPLOAESDQAAEEBAKDEGE-----EKQKEPTKSPSPSPVNSNETTSFPKKFFTH 290
 DB 3265 KPVKLTIV-SDSIPPEPKPKDSQIPSVLPDITTKPKTKTKTKPTKTEDTDQQVPDEPTTV 3323
 QY 291 GWAGWRKTSFKSKSDEDDLETAERKEQEAQKVDDEB-----KETEPASESQEP 340
 DB 3324 DTTDIPELTPTQTAQPEDTATAGITPSAQBEKSTQDDTKDTIQTKVHKHTKP--DTQKS 3381
 QY 341 AEDTDQARLSADY-----EKVELPLEDOVGDLEASSEKCAPLATEVFDE--- 385
 DB 3382 VETSELPEVHKDYQISIIHBEELVEEQEKILEVRVID-EVAEVEESQPIVEEVEDSEPQ 3440
 QY 386 -----KMEAHQBVAEVHVSTVEKTEEEQGGGGEAGGVVVE 422
 DB 3441 PATEETVEDVTPKSKKKKVVKKKTDHDELIKMLQEQIEKTELEKYEIEFDVPVKLK 3500
 QY 423 GTGESLPPEKLARPQEV-----QEAEPAEEL-----MKSR 453
 DB 3501 PEFAALEPIKIERKEQKPTKVTILDATVPKTVKLPKSKRKEKPAEELTVQLPKFRLKAR 3560
 QY 454 EMCVSGGDHTQLDLSPEKTLPKHPE-GIVSEVEMLS-----SQERIKVQGSPLKKLFS 507
 DB 3561 MVLVE-----YPPAPLIPTKTDIGAIONGELSRIEAEELTKPKPHKTKKI-- 3608
 QY 508 SSGLKSLGKKQKGGGGDEBPGEYQHITHSPESADQKGESSASSPEEPETTCLE 567
 DB 3609 -----KKIK-----DDLEKVELEKYEKYSISSEEPPEKTPYK 3640
 QY 568 KGPLEAPQDGEABEGT---TSDGEEK-REGITPWASPKMVTTPKKRVRPESDKE---- 619
 DB 3641 K-PEKAPKPEEKQEDVKLGKGGKPKKEEAP-----ENVTLKNIQKQKEVEEVELK 3694
 QY 620 ---EELEKVKSATLSSTDSTV-----SE-----MDEKVTVGEQKPEEPKRRVDT 663
 DB 3695 QKPEVEIIVEEQTKPKDGEFVVPEFPEPSEFDRPEYVPDELEQIEHPPEIPEKVKPKSKT- 3753
 QY 664 VSWEALICVSSKKRARKASSDDEGGPRTLGSDSHRAEASKDKAAGTDAVPASTQEQD 723
 DB 3754 -----KVKPKDKSKSEPETIVSIVAGVPKEEAIPEQDVKFRKPRDARPEET 3801
 QY 724 QAQSSSPAGSPSGEGEYSTWESFKRLVTPRKKSLEKAEADSSVE-----QLS 776
 DB 3802 DSEIKLRVPFOAKDEN-----PDEQALVTP--KABEPIQEIETDKAIDDEKKPKKSKP 3853
 QY 777 TEIPESSREESWVSIKFIPORRKKRADKQEQATVEDSGPVEINE-DDPNV-----PAVVP 831
 DB 3854 KKVQPKQE-----IAKEEP--BEFEVSVKEEALVDK--PIEIEKPKDVVKVKKKPEAP 3905
 QY 832 LSEYNAVEREKMEAGNTELP---OLLGAVVYVEELSCTLVHTVSVAVIDGTRAVTS--V 886
 DB 3906 VSEVVVIEEPEKPEEVEEIPVEYKITTTLVLEPEDAPKE--HQVKVIDFQERQTTTEVI 3963
 QY 887 EER-----SPSWISASVTEPLEH-----TAGEMPVPEVEVTEKDIJAETPVL 929
 DB 3964 EKVWTKKKPKQKQPEPEFEVTLKEPKEEQIQPDVVSABEISLPIEPEQKPEQVEVELKI 4023
 QY 930 TQTLPEKQAHDMVTSEVDFTSALTATSEALTTEVTEASGAEETDMVSQVAVSQ-- 987
 DB 4024 TQTTPE--EPNDVQIAVKEKVKTKPKVKVEDKIVVVEAEKQPVETIVEVEKQBEKK 4081
 QY 988 LTDSPTTE-EATPVQEVESGLDTEEE--ERQTAILOAVADK-----VKE----- 1031
 DB 4082 KSKPKSYEPKISETOSIEEKPVEVAAEAPETPKVVEKKVAEKFDSEYFTLTKDEKV 4141
 QY 1032 -----ESQVATQTVQRTGSKALEKVEB-----VEEDSEVLASEKEDVMKPGP 1075
 DB 4142 ITVDDQPEEAP--VEVVFKKPKPEPAVEAEFVWTEPKIVEETSVETAIKQKTKPKPKD 4200
 QY 1076 VQAGAEHLAQSGETQATPESILEVPEVTADVDHATCOVIVKLOQLMEQAVAPESSELT 1135

```

Db 4201 EEBQOLAIVVESEAPVAEVEVSEAPF-----SKIYE-----BEVIAEKEKKEFT 4245
QY 1136 ----DSEITNGSTPLADSDTADGTQODET-----IDSQDSKATAVRO----- 1173
Db 4246 IRVSESEBPEEPSEVGEQTVKKRKPVTFADEBPATEIIVIKESKPAEAVVTEDAHIKTKKPK 4305
QY 1174 SQVTEEBEATAQKEPSTLPNNVPAQEHGSEBGRDVLPTQOELTAAVPLAKT-EVG 1232
Db 4306 KKVTDVEAEELKIKITTEBPOEIPLEEVSEB--EVITETIK--TAPVVEEKTYSKIG 4358
QY 1233 QEGEVDWIDGKEVKEEQQVFVHSGPNSQKADAVTYDSEVMGVAGCQEKESTEVQSLSEB 1292
Db 4359 IKETPEKPAEAIVEEBPVVTEPIEBAPKEVFEHKKVRI-----EETPRELVE 4409
QY 1293 GEMETDVEKEKRETKPEQVSEBGEQETAPEHEGTYGKPVLTLMPSSEBKGALGSLGS 1352
Db 4410 EVIEEBEVKVIKRRKKPKPEIKEPPEAEVTV-----STPKPV-----EVEATSSIAVI 4456
QY 1353 PSLPDQDAGCIEVQVQSLDTVTQTAFAVEKIVETVISE-----TGESP 1398
Db 4457 PEQPTPEEBAADLKITIEEB--TPQOELVOEIEIEIEVEBKAPBEOPTDFTFATKDS 4513
QY 1399 ECVGAHLPAEKSSATGHWTLQHAEDTVPLG--PESQAESIPITVTPAPESTLHPDLQ 1455
Db 4514 KKPVTVEELPEBOV-----TIQKKKKAPVEVEEBEPAEFVVKPTVQEVTV----BE 4562
QY 1456 GEISASQRESESEEDKPDAGPDADGKESTAIKVLKABEILEL-----ESKSNKIY 1507
Db 4563 AKITSKKPKVKEE-----AAAEKLVITIEEIPTEBEVQEIIEIEIEEBKPAEYV 4614
QY 1508 LNVIOATAVDQPARTET-----APETHAVDSQ-TQVPACRLDSREPRRCWTMKDAM 1558
Db 4615 IEVKESQPEAVEDKEVSLPKKKPKAPIVEEBEPAEITLKPVKSEBQBEAKIVKKKPKKI 4674
QY 1559 -----KHPYQOPREDLOVLTVLEAWAOPKCLPR-----LOLKAP 1593
Db 4675 DEVAVADELTVKVEEVEVPEPIVEEBVIEEFIKKKPKPEBPEBIDVDAIIVKLKPP 4730

```

Search completed: December 13, 2002, 00:12:16
 Job time : 141 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:10:50 ; Search time 38 Seconds
(without alignments)
1235.763 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSTEQSPQAGSDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5421.5	67.2	1346	2	US-08-635-121-2
2	4274.5	52.9	1780	1	US-08-769-309A-5
3	4274.5	52.9	1780	3	US-08-994-570-5
4	952.5	11.8	292	1	US-08-769-309A-3
5	952.5	11.8	292	3	US-08-994-570-3
6	656	8.1	651	1	US-08-769-309A-17
7	656	8.1	651	3	US-08-994-570-17
8	639.5	7.9	453	1	US-08-769-309A-14
9	639.5	7.9	453	3	US-08-994-570-14
10	521.5	6.5	396	1	US-08-769-309A-15
11	521.5	6.5	396	3	US-08-994-570-15
12	462	5.7	1786	4	US-08-973-462-8
13	388	4.8	1805	1	US-07-853-913-2
14	365.5	4.5	3696	4	US-09-134-001C-5080
15	361.5	4.5	2101	1	US-08-466-390-4
16	361.5	4.5	2101	1	US-08-470-950-4
17	361.5	4.5	2101	1	US-08-467-781-4
18	361.5	4.5	2101	2	US-08-483-924-4
19	361.5	4.5	2101	4	US-09-452-294-1
20	361	4.5	2137	4	US-09-134-001C-4463
21	358.5	4.4	2101	1	US-08-195-487-4
22	358.5	4.4	2101	5	PCT-US93-06160-4
23	354	4.4	1618	1	US-07-853-913-4
24	354	4.4	2409	6	5180808-2
25	334.5	4.1	1162	2	US-08-728-323A-2
26	334.5	4.1	1162	4	US-09-238-568-2
27	334	4.1	1018	1	US-08-072-610-2

Query Match 67.2%; Score 5421.5; DB 2; Length 1346;
Best Local Similarity 91.6%; Pred. No. 0;

ALIGNMENTS

RESULT 1
US-08-635-121-2
; Sequence 2, Application US/08635121
; Patent No. 5910442
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,121
; FILING DATE: 19-APRIL-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/33603
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1346 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-635-121-2

28 334 4.1 1018 2 US-08-719-822B-2 Sequence 2, Appli
29 334 4.1 1018 4 US-09-092-458-2 Sequence 2, Appli
30 321 4.0 1852 1 US-08-425-061-24 Sequence 24, Appli
31 321 4.0 1852 2 US-08-825-886-24 Sequence 24, Appli
32 321 4.0 1863 1 US-08-425-061-16 Sequence 16, Appli
33 321 4.0 1863 1 US-08-480-784-2 Sequence 2, Appli
34 321 4.0 1863 1 US-08-483-553-2 Sequence 2, Appli
35 321 4.0 1863 1 US-08-487-002-2 Sequence 2, Appli
36 321 4.0 1863 1 US-08-488-011B-2 Sequence 2, Appli
37 321 4.0 1863 1 US-08-488-011B-2 Sequence 2, Appli
38 321 4.0 1863 2 US-08-825-886-16 Sequence 16, Appli
39 321 4.0 1863 4 US-08-850-727-2 Sequence 2, Appli
40 321 4.0 1863 5 PCT-US95-10203-2 Sequence 2, Appli
41 321 4.0 1863 5 PCT-US95-10203-2 Sequence 2, Appli
42 321 4.0 1863 5 PCT-US95-10203-2 Sequence 2, Appli
43 320 4.0 1898 1 US-08-056-200-94 Sequence 94, Appli
44 320 4.0 1898 2 US-08-800-644-94 Sequence 94, Appli
45 319.5 4.0 1939 4 US-09-310-187A-1 Sequence 1, Appli

Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;

```

QY 387 MEAHOEVAAVAVSTVEKTEBEEQGGGGAEGGVVEGTGESSLPREKLAEPQEVPOEAEP 446
Db 1 MEAHOEVAAVAVSTVEKTEBEEQGGGGAEGGVVEGTGESSLPREKLAEPQEVPOEAEP 60
QY 447 BELMKSRMCMVSGGDHQTLPDSPEKTLPHKHPREGIVSEVMMLSGOERIKYQSGSLKTLF 506
Db 61 BELMKSRMCMVSGGDHQTLPDSPEKTLPHKHPREGIVSEVMMLSGOERIKYQSGSLKTLF 120
QY 507 SSSGKLKLGKKOKKRGCGGDEEBEGEYOHITHESPESADQKSGSSASPEPETTC 566
Db 121 SSSGKLKLGKKOKKRGCGGDEEBEGEYOHITHESPESADQKSGSSASPEPETTC 180
QY 567 EKGPLEAPQDEABEGTTSDEKKRE---GITPMASFKNMVTPKKRVRPSESDKEEL 622
Db 181 EKGPLEAPQDEABEGTTSDEKKRE---GITPMASFKNMVTPKKRVRPSESDKEEL 234
QY 623 EKVSAATLSTDSVTSEMEQDEVKTVGEEOKPEEPKRRVDTVSWEALICVSSKKRARKA 682
Db 235 EKVSAATLSTDSVTSEMEQDEVKTVGEEOKPEEPKRRVDTVSWEALICVSSKKRARKA 294
QY 683 SSSDEGEPRTLGGDSHRAEASKDEAGTDAVPASTQEOPOAGSSAPSPAGSPSEGG 742
Db 295 SSSDIR-GPRTLGGGQSGSRGGQQRRTDAVPASTQEOPOAGSSAPSPAGSPSEGG 353
QY 743 VSTWSEFKRLVTPRKSKSKLEEK-AEDSVYQSTELIERSRESWVS--IKKPIGRRK 799
Db 354 VSTWSEFKRLVTPRKSKSKLEEKAGRTLIVYAGCPLRSNRVEKNLGFPLRNSPDGGR 413
QY 800 KRAQKQQAQAVESDGPVINEEDDNNVAVVPLSYNAVEEKMAQONTLPOLLGAVY 859
Db 414 KQWGRQQAQAVESDGPVINEEDDNNVAVVPLSYNAVEEKMAQONTLPOLLGAVY 472
QY 860 VSEELSKTLVTVSAVAVDGRATVSVESRSPWISASVTEPLETAGEAMPVEVEVEK 919
Db 473 VSEELSKTLVTVSAVAVDGRATVSVESRSPWISASVTEPLETAGEAMPVEVEVEK 532
QY 920 DIIAETPVLTLQTLPEGKADHDVMTSEVDTSEAATLSTSEALRTEVEATEASAEETT 979
Db 533 DIIAETPVLTLQTLPEGKADHDVMTSEVDTSEAATLSTSEALRTEVEATEASAEETT 592
QY 980 DMVSAVSQLTSPDPTTEATPVQEVESGVLDTSEERQTOALQAVADKVEESQVPAQTQ 1039
Db 593 DMVSAVSQLTSPDPTTEATPVQEVESGVLDTSEERQTOALQAVADKVEESQVPAQTQ 652
QY 1040 TVORSGSALEKVEBEVEDSEVLASEKEKDVMPKGPVOBAGAEHLAOGSETQATPESLE 1099
Db 653 TVORSGSALEKVEBEVEDSEVLASEKEKDVMPKGPVOBAGAEHLAOGSETQATPESLE 712
QY 1100 VPEVTADVHVATCOVILQOLMEQAVAPESSETLTDSEINGSTPLADSDTADGTQDET 1159
Db 713 VPEVTADVHVATCOVILQOLMEQAVAPESSETLTDSEINGSTPLADSDTADGTQDET 772
QY 1160 IDSQDSKATAAVROSQVTEBEAATQKEPSTLPNNVPAQEBHGEPRDVLLEPTQOELT 1219
Db 773 IDSQDSKATAAVROSQVTEBEAATQKEPSTLPNNVPAQEBHGEPRDVLLEPTQOELT 832
QY 1220 AAAAPVLAKEVGOGEVDWLDGEKVEKEOEYFVHSGNSQKADAVTADSEWVGAAGCOE 1279
Db 833 AAAAPVLAKEVGOGEVDWLDGEKVEKEOEYFVHSGNSQKADAVTADSEWVGAAGCOE 892
QY 1280 KESTIEVQSLSEEGEMETDVEKEKETKPEQVSEBEOETAPEHEGTYGKVLTLDMPS 1339
Db 893 KESTIEVQSLSEEGEMETDVEKEKETKPEQVSEBEOETAPEHEGTYGKVLTLDMPS 952
QY 1340 SERGKALGSLGSPSLPPQDKAGCTIEVQVSLDTTDTVTAQAVKVIETVVISSETGESPE 1399
Db 953 SERGKALGSLGSPSLPPQDKAGCTIEVQVSLDTTDTVTAQAVKVIETVVISSETGESPE 1012
QY 1400 CVGAHLPLPAEKSSATGGMWTLQHAEDTVPLGPEQAESPIITVTPAPSTLHPDLQGETS 1459
Db 1013 CVGAHLPLPAEKSSATGGMWTLQHAEDTVPLGPEQAESPIITVTPAPSTLHPDLQGETS 1072

```

```

QY 1460 ASQRESEEDKPDAGDADGKESATKYLKAPETLLEESKSNKTVLVNVIQTAVNOQA 1519
Db 1073 ASQRESEEDKPDAGDADGKESATKYLKAPETLLEESKSNKTVLVNVIQTAVNOQA 1132
QY 1520 RTEPAPETHAVDSOTQVPACRLDSREBNRCWTMKDKAMKHPVQPREDDQVLTLEBAW 1578
Db 1133 RTEPAPETHAVDSOTQVPACRLDSREBNRCWTMKDKAMKHPVQPREDDQVLTLEBAW 1191

RESULT 2
US-08-769-309A-5
; Sequence 5, Application US/08769309A
; Patent No. 5741890
; GENERAL INFORMATION:
; APPLICANT: Scott, John D.,
; APPLICANT: Nauert, Brian J.,
; APPLICANT: Klaus, Theresa M.
; TITLE OF INVENTION: Protein Binding Domains of Gravin
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,309A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5741890and, Greca E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1780 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-769-309A-5

Query Match 52.9%; Score 4274.5; DB 1; Length 1780;
Best Local Similarity 57.8%; Pred. No. 9,2e-255;
Matches 957; Conservative 187; Mismatches 402; Indels 109; Gaps 36;

```

Db 235 TEKPEETLKREQSHAEISPPAESQAV--BECKEKEGEKQEKPKSAESPTSPVTSGETS 293
Qy 283 SFKKPFTHWAGWKKTSFKKSKEDDLLETABKRKEQEAKEDEBEKTEKPASBE-----337
Db 294 TFKKFFTOGWAGWKKTSFRKPKEDVEASESKKQEPKVDTEEDGKAEVASEKLTASE 353
Qy 338 ---OEPAEDTOARLSADYKVLPLPDQVDGLAEASSEKCAPLATEVDFEKMEHQ-E 392
Db 354 QAHPQEPAESAHPRLSAEYKVLPEEQVSGQSPSEKPPAPLATEVDFEKEVHQEE 413
Qy 393 VVAEYHVSTVEKTEBEEOGGGAEAGVVEGTGSLPEKLAEPQVQBAEPQBAEELMKS 452
Db 414 VVAEYHVSTVEETEEO-----KTEVEETAGSVPAEELVGMDAEPQBAEPKAKELVKL 465
Qy 453 REMCVSGDHTQLTDLSPKEETLKPHEGIVSEVEMLSQBRIRKVGQSPKLKLFSSGLK 512
Db 466 KETCVSGEDPTQADLSPDEKVLSPKPEGVVSEVEMLSQBRIRKVGQSPKLKLFSTGLK 525
Qy 513 KLSGKKKGRGGGDBEPGEGYQHHTESPSADEQKGESSASSPPEETTCLEKPLE 572
Db 526 KLSGKKKGRK--GGDBESGHTQVPADSPDSQBEQKGESSASSPPEETTCLEKGLAE 584
Qy 573 APODGEAEEGTSDGKKREGITPWASFKMVTPKKVRPSESDEKBELEKVKSATLSS 632
Db 585 VQDGEAEAGTSDGKKREGITPWASFKMVTPKKVRPSESDEKDELKVKSATLSS 644
Qy 633 TDSTVSEMDQEVKTVGBEQKPEEKRRVDTSVSWEALICVGSKKRKARAKSSDDEGGPR 692
Db 645 TESTASEMEQEMKGSVEEPKPEEKRRVDTSVSWEALICVGSKKRRARRSSDEEGPK 704
Qy 693 TLGDSHRAEASXDKAGTDVAPASTQEOQOAGSSPEPAGSPSGEGVSTWESPKRL 752
Db 705 AMGGDHOKADAGKDKETGDTGILAGSQEHDPGQGSSSPEQAGSPTEGEGVSTWESPKRL 764
Qy 753 VTPRKSKSLEEKAE-----SSVEQLSTETEPSRESWSIKFIPCKRRKRDAGKQEQ 808
Db 765 VTPRKSKSLEEKESDIAGSGVHESTPDTEPKGESSWSIKFIPGRRKKRPDQKQEQ 824
Qy 809 ATVEDSGPVEINDDPNVPAVPLSEYNAVERKME---AQNTPELQQLGAVVYSBELS 865
Db 825 APVEDAGPTGANEDSDVPAVPLSEYDAVERKMEQAQKGAEPQEKAAETVSKELS 884
Qy 866 KTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPVVEVTEKDIIA-E 924
Db 885 ESQVHMMAAAVADGTRAATIIERSPSWISASVTEPLEQVEAEAAALLTEVLEREVIAGE 944
Qy 925 ETPVLTOTLPEGKAHDMVTSEVDFTEAVTATETSEALRTEETVTEASGAEETDMVSA 984
Db 945 EPPTVTPLPENRARGDVTVSEALTEPAVTAETAAGPLGSEEGTEASAAETTEMVSA 1004
Qy 985 VSQLTSDPTTEATPVQESVGLDTEEBERQTOAILQAVADKVKESQVPAE-----1038
Db 1005 VSQLTSDPTTEATPVQESVGLDTEEBERQTOAILQAVADKVKESQVPAE-----1064
Qy 1039 -QTVORTGSKALEKVEVEEDSEVLASEKEKDVMPKGPVQAGAEHLAQGSETQCATPES 1097
Db 1065 LQPVQRA-----BAERPEQAEASGLKETTDLVVLKVDQAQEAKEPTTQGVVGTTPES 1118
Qy 1098 LE-VPEVTADVH---VATQV-----IKLOOL-MEQAVAPESSETTDSNKGSTPLAD 1147
Db 1119 FEKAPQVTESESSELVTTQCAETLAGVKSQEMVMEQAIIPDSVETPTDSETDGSSTPAD 1178
Qy 1148 SDTADGTQOEDTISQDSKATAAARQSVQVTEEAATAQKEPSTLPNNVPAQEBHREPG 1207
Db 1179 FDAGTTQKDBIVIHENEVHLV-PVRGTBAEAVPAQEKERPPA-PSFVFPQEBTKQSK 1236
Qy 1208 -RDVLEPTQOELTAAAVPLAKTEVGEVGEVDWLDGKVK-----ERQEVFVHSG--PNS 1259
Db 1237 MEDTLEHTDKREVSVEVTSILSKTEGTQ--EADQVADKTKDVPFEGLEGSIDGITVSR 1294
Qy 1260 QKAADVITYDSEVMGACQEKESTEVQS-----LSLEGEEMETDVEKEKTEKPEQVSEE 1314

Db 1295 EKYTEVALKCEGTEEAECCKDDALELOSHAKSPSPVEREMVQVBERKTEAPETHNEE 1354
Qy 1315 G-EOETAAPHEGTYGKPVLTLDMPSSERKALGSLGGS-PSLPDQDKAGCIEVQVQSLD 1372
Db 1355 KLEHETAVTVSEEVSKQLLTQVNVPIIDGAKEVSSLESGSPPCLGQBEAVCTKIQVQSS 1414
Qy 1373 TTVTQTAEAVEKVI--ETVVISSETGSPCVGAHLLPAEKSSATGGHWTLOHAEDTVPGL 1430
Db 1415 ASFTLTAAABEEKVLGTETANILETGETLEPAGAHVLEEKSSEKKNEDFAAHPCGEDAVPTG 1474
Qy 1431 PESOAEIPIIIVTPAPESTLHPDLQGEISASQERSERSEEDKPDAGPDADGKESFAIEKVL 1490
Db 1475 PDQAKSTPVIIVATYKGLSSDLGEKTTSLKWSKDEVDQVACQEV--KVSVAIEDL- 1531
Qy 1491 KAPPE--ILELESKNKIVLNVITQAVDQPART-ETAPETHAYDSOTQVPCACRLDSREP 1547
Db 1532 --BPENGILSELTSSKLVQNILOTAVDQFVRTEETATEMLTSELQTOAHVIRADSQDAG 1589
Qy 1548 RCWTK-----MKDAKMKHPVPQPREDLQVLTVLEA 1577
Db 1590 QETEKEGEPQASAOETPITSKESESESTAVGQA 1624

RESULT 3
US-08-994-570-5
; Sequence 5, Application US/08994570
; Patent No. 6090929
; GENERAL INFORMATION:
; APPLICANT: Scott, John D.,
; APPLICANT: Nauert, Brian J.,
; APPLICANT: Klauack, Theresa M.,
; TITLE OF INVENTION: Protein Binding Domains of Gravin
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994,570
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090929and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1780 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-994-570-5

Query Match 52.9%; Score 4274.5; DB 3; Length 1780;
Best Local Similarity 57.8%; Pred. No. 9.2e-255;
Matches 957; Conservative 187; Mismatches 402; Indels 109; Gaps 36;
Qy 1 MGAGSSTEQRSPEQ-PAGSDTSELVLSGHGPAEAS-GAAGDPADA--DPATKLPKNG 56
Db 1 MGAGSSTEQRSPEQPPGSGSTPAEPSPGSGSAEAAADTTADPAIAASDPATKLPKNG 60

QY 57 QLSVNGVABEGDVHVOENOBQ-----EEVVDVEDVGORSEDEKREKOR 102
 Db 61 QLSVNGVABEGDELIOGDLNGOGALNGBEVEEIVTEVQORSEDEKREKOR 120
 QY 103 VEEAANSTAAVEDITKQOQOETSEIIEOI PASENNVEEMVOPAESQADVGFKKVFKVVG 162
 Db 121 DKEAATKSAVVDITDDQOEENRN-IEOI PSESNLVEELTQPTESQANDIGKKVFKVVG 179
 QY 163 FKFTVKDKNEKSDTVOLLTVKKDGEAGASVAGDHQEPBSVETAVGSASEKSELKOS 222
 Db 180 FKFTVKDKNEKSDTVOLLTVKKDGEAGASVAGDHQEPBSVETAVGSASEKSELKOS 234
 QY 223 TEKQOBTIKQOOSSTEIPLQASDQAABEADBEKEQEKPTSPSPSPVNSSTTS 282
 Db 225 TEKQOBTIKQOOSSTEIPLQASDQAABEADBEKEQEKPTSPSPSPVNSSTTS 293
 QY 283 SPKFFFTGMAGMKRKTSPFKSKEDDLTAETAKRQOAKVDEEKEKTEPASEE----- 337
 Db 294 TPKKFTGMAGMKRKTSPFKSKEDDLTAETAKRQOAKVDEEKEKTEPASEE----- 353
 QY 338 ----OEPAEDTDQARLSADYKVELPLBEDQVGLBASSEKCAPLATEVDEKMAHQ-E 392
 Db 354 QAHPEPAPASNEPRLSAEYKVELPSEBQVSGSQSPSEKAPLATEVDEKMAHQ-E 413
 QY 393 VVAEYHVSVEKTEBEGGSGGAEAGVYVEGTEGSLPPEKLAEPQVQVQEAAPAEELMKS 452
 Db 414 VVAEYHVSVEKTEBEGGSGGAEAGVYVEGTEGSLPPEKLAEPQVQVQEAAPAEELMKS 465
 QY 453 REMCVSGDHTDLSPEKTLPRHPEGIVSEVEMLSQERIKYQGSPLKLFSSGLK 512
 Db 466 KETCVSGEDPTQADLSPEKTLSPREKVLSPREGVSEVENTLSQERIKYQGSPLKLFSSGLK 525
 QY 513 KLSGKKOKRGKGGEDEBGEYQIHTESPESADEOKGSSASSPPEPETTCLEKPLE 572
 Db 526 KLSGKKOKRGKGGEDEBGEYQIHTESPESADEOKGSSASSPPEPETTCLEKPLE 584
 QY 573 APQDEABEERTSDDEKREGITPMASPKMVTTPKKRVRPESPEDEKLEKVSATLSS 632
 Db 585 VQODEABEERTSDDEKREGITPMASPKMVTTPKKRVRPESPEDEKLEKVSATLSS 644
 QY 633 TDSVTSEMODEKTVGEOBKPEPKRVRVTSWEALICVSSKKRARRASSDDEGGR 692
 Db 645 TESTISEMOEEMKGVSEBKPEPKRVRVTSWEALICVSSKKRARRASSDDEGGR 704
 QY 693 TLGDSHRAEASKDEAGTDAVPASTOEDQOAGSSSPBAGSPBEGGVSTWESFKXL 752
 Db 705 AMGSHQADEAGKDEAGTDAVPASTOEDQOAGSSSPBAGSPBEGGVSTWESFKXL 764
 QY 753 VTPRKKSXSLEKKAED----SSVEQLSTEIPSPRESVNSIKKTPGRRKKAQOQO 808
 Db 765 VTPRKKSXSLEKKAED----SSVEQLSTEIPSPRESVNSIKKTPGRRKKAQOQO 824
 QY 809 ATVEDESGPEINEDDPNVAVVPLSEYNAVVEBKME---ACQNTLPLGLAVVYSELS 865
 Db 825 AVEDESGPEINEDDPNVAVVPLSEYNAVVEBKME---ACQNTLPLGLAVVYSELS 884
 QY 866 KTLVTVSAVVIDGTAVTVSEERSPSWISASVTEPLEHTAGEAMPVEVTEKQIIA-E 924
 Db 885 ESQVHMAAAVADGTAAITIEERSPSWISASVTEPLEHTAGEAMPVEVTEKQIIA-E 944
 QY 925 ETPVLTQTLPEGKDAHDMVTSVDFTSAAVATETSEALRTBEVYEAAGAETTDMVA 984
 Db 945 ETPVLTQTLPEGKDAHDMVTSVDFTSAAVATETSEALRTBEVYEAAGAETTDMVA 1004
 QY 985 VSQULDSPTTEAPTPVQVESGVLTDEEERQTOAILAVVADKVEEEOVAT----- 1038
 Db 1005 VSQULDSPTTEAPTPVQVESGVLTDEEERQTOAILAVVADKVEEEOVAT----- 1064
 QY 1039 -QTVORTGSKALEKVEEVEDESVLASEKXKDVMPKGVQVQEAAGHLAGSETQATPES 1097
 Db 1065 LQPVQRA-----EAHREPEQABASGLKETTVLVKVDQAEKTEPFTQGVKVGQTTSES 1118
 QY 1098 LE-VPEVTADVDH---VATQV-----IKLQOL-MEQAVAPESSEFTLTSSTNGSTPLAD 1147

Db 1119 PEKAPQVTESTIESSELVTTQAEETLAGVKSQEMMEQAIIPDVSJETITDSTOSTVAD 1178
 QY 1148 SDTADGTQOEDTIDSDSKATPAVQSOVTEEBAAVQKEEPSTLPNNVPAOEHEGEPG 1207
 Db 1179 FDAGTQOKEIIVIEHENEVHLV-PVRGTEAEAVPAQKERPPA-PSSFVQOETTKQSK 1236
 QY 1208 -RDVLEPQOELTAAVPLVAKTVEQEGVDMIDGKKV-----EQEYFVHSG--PNS 1259
 Db 1237 MEDTLEHTDKESVSETVSLSTKTEGTQ--EADQYADEKTDVPPFEGLEGSIDGITVSR 1294
 QY 1260 QKADVTYDSEFVMVAGQOKEESTEVOS-----LSLEEGEMETVEKEKETKPEOYSEE 1314
 Db 1295 EKVTEVALKGBGTAEACKODDALELOSHAKSPSPYERENAVVVEKEKTEAEPTHNEE 1354
 QY 1315 G-EQETAAPEHGTGYKPVLTLDMPSSERKALGSLGS--PSLPDQKAGCIEVOVQSLD 1372
 Db 1355 KLHETAVTVSEVSKQLQVNVNPIIDGAKEVSSLGSPPPCLGQEAHVCTKIQVQSE 1414
 QY 1373 TTVVQTAFAVKVI--ETVVISSEGESPEVCYGAHLPAEKSSATGGMWTLQHADTVPLG 1430
 Db 1415 ASFTLTAABEKEKVLGETANTLETGETLEPAGAHVLEEKSSXENEDFAHPGEDAVPTG 1474
 QY 1431 PESQAESIPITVPAPSTLHPDLQGERISAGQERSEEDKPDADGKESTAEIKVL 1490
 Db 1475 PDQAKSTPVIVSATTKKGLSSDLEGEKTTSLKMKSDVDEQVACQEV--KVSVAIEDL- 1531
 QY 1491 KAEBE--LLEESKSNKIVLVNVIQTAVDQFART-ETAPETHAYDSQTPVAPCRDLSREP 1547
 Db 1532 --EPENGILELETSSKLVQNIIDTAVDQFVRTETETEMLTSELQTAHVIKADSDQAG 1589
 QY 1548 RCWTK-----MKDAMKHPVPQREDLOVLTVLEA 1577
 Db 1590 QETKEGEEPOASQADETPITSAKESSESTAVGQA 1624

RESULT 4
 US-08-769-309A-3
 Sequence 3, Application US/08769309A

GENERAL INFORMATION:
 APPLICANT: Scott, John D.,
 APPLICANT: Nauert, Brian J.,
 APPLICANT: Klauack, Theresa M.,
 TITLE OF INVENTION: Protein Binding Domains of Gravin
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower/233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/769,309A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5741890and, Grete E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/33451
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-474-6300
 TELEFAX: 312-474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 292 amino acids

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-769-309A-3

Query Match      11.8%; Score 952.5; DB 1; Length 292;
Best Local Similarity 66.1%; Pred. No. 2.5e-51;
Matches 199; Conservative 30; Mismatches 53; Indels 19; Gaps 4;

QY 254 KDSGEKQKEPTKSPSPVNSSETTSFKKFTTHGAGWRKKTSTFKSKEDDLTAE 313
Db 1 KEEGEKQKEPSKSAESPTSPVTSETGTFKKFTTQGWAGWRKKTSTFRKPKEDVEASE 60
QY 314 KRKEQAEKVDEEKEKTEPASEE-----QEPADTDQARLSADYKVELPLEDOV 364
Db 61 KKKEQPEKVDTEEDGKAESAELTASEQAHQPEPAESAHEPRLSAEYKVELPSEEQ 120
QY 365 GDLEASSEEKCAPLATEVFDEKMEAHQ-EVVAEVHVSTVEKTEEQGGGGEAGGVVVEG 423
Db 121 SGSGPSEKPAFLATEVFDEKIEVHQEEVVAEVHVSTVEERTEEQ-----KTEVEE 172
QY 424 TGBSLPPEKLAEPQEVQPAEAPAEELMKSRMCMVSGGDHTQLTDLSPPEKTLPKHPGIV 483
Db 173 TAGSVPAEELVGMDAEPQAEPAKELVKLKTCTVSGEDPTQGDLSPEKVLSPPEGVV 232
QY 484 SEVEMLSQERIKVQGSPLKLFSSGLKLSGKKQKGGGGDEBPGEYQHIHTESPE 543
Db 233 SEVEMLSQERIMVQGSPLKLFSTGLKLSGKKQKGR-KGGDEBSGEBHTQVPADSPD 291
QY 544 S 544
Db 292 S 292
```

```
RESULT 5
US-08-994-570-3
; Sequence 3, Application US/08994570
; Patent No. 6090929
; GENERAL INFORMATION:
; APPLICANT: Scott, John D.,
; APPLICANT: Nauert, Brian J.,
; APPLICANT: Klauck, Theresa M.
; TITLE OF INVENTION: Protein Binding Domains of Gravin
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994,570
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090929and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; US-08-994-570-3

Query Match      11.8%; Score 952.5; DB 3; Length 292;
Best Local Similarity 66.1%; Pred. No. 2.5e-51;
Matches 199; Conservative 30; Mismatches 53; Indels 19; Gaps 4;

QY 254 KDSGEKQKEPTKSPSPVNSSETTSFKKFTTHGAGWRKKTSTFKSKEDDLTAE 313
Db 1 KEEGEKQKEPSKSAESPTSPVTSETGTFKKFTTQGWAGWRKKTSTFRKPKEDVEASE 60
QY 314 KRKEQAEKVDEEKEKTEPASEE-----QEPADTDQARLSADYKVELPLEDOV 364
Db 61 KKKEQPEKVDTEEDGKAESAELTASEQAHQPEPAESAHEPRLSAEYKVELPSEEQ 120
QY 365 GDLEASSEEKCAPLATEVFDEKMEAHQ-EVVAEVHVSTVEKTEEQGGGGEAGGVVVEG 423
Db 121 SGSGPSEKPAFLATEVFDEKIEVHQEEVVAEVHVSTVEERTEEQ-----KTEVEE 172
QY 424 TGBSLPPEKLAEPQEVQPAEAPAEELMKSRMCMVSGGDHTQLTDLSPPEKTLPKHPGIV 483
Db 173 TAGSVPAEELVGMDAEPQAEPAKELVKLKTCTVSGEDPTQGDLSPEKVLSPPEGVV 232
QY 484 SEVEMLSQERIKVQGSPLKLFSSGLKLSGKKQKGGGGDEBPGEYQHIHTESPE 543
Db 233 SEVEMLSQERIMVQGSPLKLFSTGLKLSGKKQKGR-KGGDEBSGEBHTQVPADSPD 291
QY 544 S 544
Db 292 S 292
```

```
RESULT 6
US-08-769-309A-17
; Sequence 17, Application US/08769309A
; Patent No. 5741890
; GENERAL INFORMATION:
; APPLICANT: Scott, John D.,
; APPLICANT: Nauert, Brian J.,
; APPLICANT: Klauck, Theresa M.
; TITLE OF INVENTION: Protein Binding Domains of Gravin
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,309A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5741890and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-769-309A-17
```


Query Match	8.1%;	Score 656;	DB 1;	Length 651;
Best Local Similarity	38.3%;	Pred. No. 1.4e-32;		
Matches 191;	Conservative 80;	Mismatches 188;	Indels 40;	Gaps 177;

Qy	1110	VANVOV-----IKLOOL-MEOAAVAPSSSEFLTPSENGSTPLADSDITADGTOEDTISQ	1163
Db	6	VTTQOAEFTLAGVSGQWEMEQALPPROSVETPIPDSETIDSGTVPADDPACETQYKDEIVIH	65
Qy	1164	DSKATAVROSQVTEEEAATAQKEBESTLPNNVPAOEHEGSEPG-RDVLEPTQDELTAA	1222
Db	66	EENEVHLY-PYRGTEAEVAAYAKERPPA-PSSVPQOEETKEQSKMEDTLEHTDKESVET	123
Qy	1223	VPLYAKTEVGOEGEVDMLDEKVK----EBOEVFVHSG-PNSQKADVTYDSVMGVA	1275
Db	124	VSLISTEGTQ--EADQYADEKTKDVPFEGLEGISIDTOITYSREKTEVETALKGTEEA	181
Qy	1276	GCOKESTEVOS-----LSIEBEGMEDVKEKRETKPQVSEBEG-EQPTALPHEHYTG	1329
Db	182	ECKDDALELEIASHKSPPSFVEREMVQVREKTEAEPTHVNEEKLEHETAVTYSEEVSK	241
Qy	1330	KPVLTLMPSSEKKGALSGSGS-PSLPDDKAGCIEVQVOSIDTVPQTAAEVKVI-	1386
Db	242	QLQTVAVPPIIDGAKEVSSLEGSPPCLOGEEAVCTKIOVQSEKASFTLPAABEEKVLG	301
Qy	1387	ETVVISSESGPECVGAHLLPAEKSSATGKMTLQHAEDTVPPLGSPESQAESIPILVTPAP	1446
Db	302	ETANILETGETLEPAGHVLVBEKSSKEMEDPAHNGEDAVPRPGCOAKSPVIVASMTT	361
Qy	1447	ESTLHPDLOGEISASOBERESEEDKPDAGBDADGKESTALEKYLKAEF--ILELESKN	1504
Db	362	KKGLSSDLEGEKTTSLMKKSDDEVQVACGEV--KVSVAIEDP---EPENGILELETKSS	416
Qy	1505	KIVLVNIOTAVDQFART-ETAPETHAVDSQTVQFACRLDSRENRCKWK-----WKDAM	1556
Db	417	KLVONIIOITAVDQFVRETEATATEMLTSELQOAHVIKADSQDAGQETEKERGEBPQASAD	476
Qy	1559	KHPVQPREDLQVLTVLEA 1577	
Db	477	ETPLTSAEESSESTAVGQA 495	

RESULT 7
 US-08-994-570-17
 Sequence 17, Application US/0894570
 Patent No. 6090529
 GENERAL INFORMATION:
 APPLICANT: Scott, John D.,
 APPLICANT: Nauer, Brian J.,
 APPLICANT: Klauk, Theresa M.
 TITLE OF INVENTION: Protein Binding Domains of Gravin
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower/233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/994,570
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6090929and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/33451
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-994-570-17

Query Match	8.1%;	Score 656;	DB 3;	Length 651;
Best Local Similarity	38.3%;	Pred. NO. 1.4e-32;		
Matches 191;	Conservative 80;	Mismatches 188;	Indels 40;	Gaps 17;

```

0Y 1110 VATQV-----IKLOOL-MEONAPRSPSTLJDSFNSTSLPLASDPADGQOQDETIDQ 1163
Db 6 VTTQAEFTLGVGSQOMVMEQALPPBSVETPTDSTSDTSPADPADGTTQKQEIYEIH 65
0Y 1164 DSKATAVRQSYTEEAATTAQKEEDSTLNNVPAQOEHEGEPG-RVLEPTQOELTAA 1222
Db 66 ENEVHVLV-PVRGTAEANPAQKEPRP-PSSFVQOEETKQSKMEDTLETTDXKEVSJET 123
0Y 1223 VPVLAKTEVQCEGEVMDLGEKVK-----EBOEVVHSQ--PNSQAKADVTYDSEVGV 1275
Db 124 VSLSTKEGQ--EAOQYADEKTKDVPFEGHSGSIDGITVSRKXTEVALKQGTBEA 181
0Y 1276 GCQKESTEVOS-----LSLEGEEMETDVEKRETKETKEQVSESG-EOETAPHEGTG 1329
Db 182 ECKDDDALEETQSHAKSPSPVEREMVQVEREKTEAPTHNEEKLHETRTAVTSEEVSK 241
0Y 1330 KPVLTLMPSSEKGLMGLSGS-PSLPRQDAGCIEVYOSLDITVQTAFAVEKV- 1386
Db 242 QLTQTVANPPIIDGAKESVLSGSPPCLOQEBEAVCTKIOVOSSEKSFLLTAAEBEKVLG 301
0Y 1387 ETVVISETGESPEVCGAHLLPAEKSSATGHHLTLOAEDTVLGPESQAESIPPIVTPAP 1446
Db 302 ETANILEGTGLBPAGHVLVLEKSSKKNEDFAHGHGEDAVPTGDCQAKSPVIVASTT 361
0Y 1447 ESTHAPDLQGEISAQORESEEDKDPADGDADGKESTAIKVLKAPBP--LLELESKN 1504
Db 362 KKGSSSDLEGEKTTSLMKKSDDEVQVQAOEV--KVSVAIEPL---EPENGILLETAKSS 416
0Y 1505 KIYLVNQTANDQAPR-ETAPETHAVDSQOTQVPAQRILDSRPNNCWK-----MKDKM 1558
Db 417 KLVNIIQTADQVFRTEETATMLTSELQTAHVIKADSDQAGETEKEGGEPOASQOD 476
0Y 1559 KHPVQPREDLQVLTVEA 1577
Db 477 ETPITSAKESESTAVQQA 495

```

```

1      RESULT 8
2      US-08-769-309A-14
3      : Sequence 14, Application US/08769309A
4      : Patent No. 5741890
5      : GENERAL INFORMATION:
6      : APPLICANT: Scott, John D.,
7      : APPLICANT: Nauert, Brian J.,
8      : APPLICANT: Klaufek, Theresa M.
9      : TITLE OF INVENTION: Protein Binding Domains of Gravin
10     :
11     : NUMBER OF SEQUENCES: 24
12     : CORRESPONDENCE ADDRESS:
13     : ADDRESSSES: Marshall, O'Toole, Gerstein, Murray & Borrun
14     : STREET: 6300 Sears Tower/233 South Wacker Drive
15     :
16     : CITY: Chicago
17     : STATE: Illinois
18     : COUNTRY: United States of America
19     : ZIP: 60606-6402
20     : COMPUTER READABLE FORM:
21     : MEDIUM TYPE: Floppy disk
22     : COMPUTER: IBM PC compatible
23     : OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,309A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5741890and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-769-309A-14

```

```

Query Match 7.9%; Score 639.5; DB 1; Length 453;
Best Local Similarity 40.5%; Pred. No. 8.7e-32;
Matches 183; Conservative 69; Mismatches 165; Indels 35; Gaps 16;

```

```

QY 1110 VATCOV-----IKLOQL-MEQAVAPESSETLTDSETNGSTPLADSDTADGTQOQDETIDSQ 1163
Db 6 VTTCOAETLAGVKSQEMVMEQAIPDPSTVETDSETDGTVPVADFDAGPTTQKDEIVEIH 65

QY 1164 DSKATAAVRQSVQVTEEAATAQKEEPTLNNVPAQEEHGEPEG-RDVLEPTQOELTAAA 1222
Db 66 ENEVHLV-PVRGTEAEAVPAQKERPPA-PSSFVFQETKEQSKMEDTLEHTDKESVET 123

QY 1223 VPVLAKEVQGEVDWLDEKVK-----EEQEVFVHSG--PNSQKAADVTYDSEVMGVA 1275
Db 124 VSILSKTEGTQ--EADQYADEKTKDVFPFEGLEGGIDTGITVSRKVTVALKGEGETEA 181

QY 1276 GCOEKESTEVQS-----LSLEGEMETDVEKEKRETKPQOVSEEG-EQETAAPPEHGTG 1329
Db 182 ECKKDDALELQSHAKSPSPVEREMVQVVEREKTEAEPTHVNEEKLEHETAVTVSEVSK 241

QY 1330 KPVLTLDMPSSERKALGSLGGS--PSLPDQDKAGCIEVQVQSLDITVTTQTAAEAVEKVI-- 1386
Db 242 QLLQTVNVPIDGAKESVSLGSPPCLGQEEAVCTKIQQVSEASFTLTAABEEKVLG 301

QY 1387 ETVVISGTSPESPCVGAHLPAEKSSATGGHWTLQHAEDTVPLGPESQAESIPITVPAP 1446
Db 302 ETANILETGTLEPAGAHVLVEKSSKNEKEDFAAHPCGEDAVPTGPDCAKSTPVIYSATT 361

QY 1447 ESTLHPDLOGEISASQRESEEDKPDAGPDAGKESHTAEIKVLKAEPE--ILELESKSN 1504
Db 362 KKGSLDLEGEKTTSLKWKSDVEVQVACQEV--KVSVAIEDL---EPENGILELETKSS 416

QY 1505 KIVLNVIOQTAVDQFART-ETAPETHAYDSQTQ 1535
Db 417 KLVQNIIQTAVDQFVRTEETATEMLTSELQTQ 448

```

```

RESULT 9
US-08-994-570-14
; Sequence 14, Application US/08994570
; Patent No. 6090929
; GENERAL INFORMATION:
; APPLICANT: Scott, John D.,
; APPLICANT: Nauert, Brian J.,
; APPLICANT: Klauck, Theresa M.
; TITLE OF INVENTION: Protein Binding Domains of Gravin
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois

```

```

; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994,570
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090929and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-994-570-14

```

```

Query Match 7.9%; Score 639.5; DB 3; Length 453;
Best Local Similarity 40.5%; Pred. No. 8.7e-32;
Matches 183; Conservative 69; Mismatches 165; Indels 35; Gaps 16;

```

```

QY 1110 VATCOV-----IKLOQL-MEQAVAPESSETLTDSETNGSTPLADSDTADGTQOQDETIDSQ 1163
Db 6 VTTCOAETLAGVKSQEMVMEQAIPDPSTVETDSETDGTVPVADFDAGPTTQKDEIVEIH 65

QY 1164 DSKATAAVRQSVQVTEEAATAQKEEPTLNNVPAQEEHGEPEG-RDVLEPTQOELTAAA 1222
Db 66 ENEVHLV-PVRGTEAEAVPAQKERPPA-PSSFVFQETKEQSKMEDTLEHTDKESVET 123

QY 1223 VPVLAKEVQGEVDWLDEKVK-----EEQEVFVHSG--PNSQKAADVTYDSEVMGVA 1275
Db 124 VSILSKTEGTQ--EADQYADEKTKDVFPFEGLEGGIDTGITVSRKVTVALKGEGETEA 181

QY 1276 GCOEKESTEVQS-----LSLEGEMETDVEKEKRETKPQOVSEEG-EQETAAPPEHGTG 1329
Db 182 ECKKDDALELQSHAKSPSPVEREMVQVVEREKTEAEPTHVNEEKLEHETAVTVSEVSK 241

QY 1330 KPVLTLDMPSSERKALGSLGGS--PSLPDQDKAGCIEVQVQSLDITVTTQTAAEAVEKVI-- 1386
Db 242 QLLQTVNVPIDGAKESVSLGSPPCLGQEEAVCTKIQQVSEASFTLTAABEEKVLG 301

QY 1387 ETVVISGTSPESPCVGAHLPAEKSSATGGHWTLQHAEDTVPLGPESQAESIPITVPAP 1446
Db 302 ETANILETGTLEPAGAHVLVEKSSKNEKEDFAAHPCGEDAVPTGPDCAKSTPVIYSATT 361

QY 1447 ESTLHPDLOGEISASQRESEEDKPDAGPDAGKESHTAEIKVLKAEPE--ILELESKSN 1504
Db 362 KKGSLDLEGEKTTSLKWKSDVEVQVACQEV--KVSVAIEDL---EPENGILELETKSS 416

QY 1505 KIVLNVIOQTAVDQFART-ETAPETHAYDSQTQ 1535
Db 417 KLVQNIIQTAVDQFVRTEETATEMLTSELQTQ 448

```

```

RESULT 10
US-08-769-309A-15
; Sequence 15, Application US/08769309A
; Patent No. 5741890
; GENERAL INFORMATION:
; APPLICANT: Scott, John D.,
; APPLICANT: Nauert, Brian J.,
; APPLICANT: Klauck, Theresa M.
; TITLE OF INVENTION: Protein Binding Domains of Gravin

```

```

NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,309A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5741890and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-769-309A-15

Query Match      6.5%; Score 521.5; DB 1; Length 396;
Best Local Similarity 38.2%; Pred. No. 1.3e-24;
Matches 147; Conservative 62; Mismatches 149; Indels 27; Gaps 12;

QY 1110 VATGV-----IKLOOL-MEOAVAPESSETLTDTSETNGSTPLADSDPTADGTQOQDETTDSQ 1163
DB 6 VTTQAEETLAGVKSQEMWMEQAIIPDSVETPTDSETDGTVPADFDPAGTQKDEIVEIH 65
QY 1164 DSKATAAROSQVTEBEATATQKEEPTLPNNVPAQEBHGEPG-RDYLEPTQOELTAA 1222
DB 66 EENEVHLV-PVRGTEAEAVPAQKERPPA-PSSFVFOEETKEOSKMEETLEHTDKEVSVE 123
QY 1223 VPLAKTEVGEQGEVDMLDGEKVK-----EEOEVFVHSG--PNSOKADVTYDSEWNGVA 1275
DB 124 VSILKTEGTQ--EADQYADEKTDVPPFEGLEGSIDTIGITVSREKVTIEVALKSGTEEA 181
QY 1276 GCOEKESTEVOS-----LSLEGEMETDVEKEKRETKPEQVSEBG-EOETAAPBHGTYG 1329
DB 182 ECKKDDALELQSHAKSPSPVEREMVQVEREKTEAEPHVNVEKLEHETAVTVESEVSK 241
QY 1330 KPVLTLDMPSSERKALSLGGS--PSLPDQKACIEVQVOSLDTTYQTAAVKEVI-- 1386
DB 242 QLOTVNVPPIIDGAKEVSSLEGSPPCLGGEAVCTKIQVOSSESFTLTAABEEKVLG 301
QY 1387 ETVVISEGESPECVGAHLLPAEKSSATGHWTLQHAEDTVPLGPESQASIPITVPAP 1446
DB 302 ETANLLETGETLEPAGALVLEEKSEKNEDEPAHNGEDAVPTGPDCAKSTPVIYSATT 361
QY 1447 ESTLHPDLQGEISASQERSEEDK 1471
DB 362 KKGSSDLEGEKTTSLKWKSDDEVDE 386

```

```

APPLICANT: Klauk, Theresa M.
TITLE OF INVENTION: Protein Binding Domains of Gravin
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,570
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090929and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-994-570-15

Query Match      6.5%; Score 521.5; DB 3; Length 396;
Best Local Similarity 38.2%; Pred. No. 1.3e-24;
Matches 147; Conservative 62; Mismatches 149; Indels 27; Gaps 12;

QY 1110 VATGV-----IKLOOL-MEOAVAPESSETLTDTSETNGSTPLADSDPTADGTQOQDETTDSQ 1163
DB 6 VTTQAEETLAGVKSQEMWMEQAIIPDSVETPTDSETDGTVPADFDPAGTQKDEIVEIH 65
QY 1164 DSKATAAROSQVTEBEATATQKEEPTLPNNVPAQEBHGEPG-RDYLEPTQOELTAA 1222
DB 66 EENEVHLV-PVRGTEAEAVPAQKERPPA-PSSFVFOEETKEOSKMEETLEHTDKEVSVE 123
QY 1223 VPLAKTEVGEQGEVDMLDGEKVK-----EEOEVFVHSG--PNSOKADVTYDSEWNGVA 1275
DB 124 VSILKTEGTQ--EADQYADEKTDVPPFEGLEGSIDTIGITVSREKVTIEVALKSGTEEA 181
QY 1276 GCOEKESTEVOS-----LSLEGEMETDVEKEKRETKPEQVSEBG-EOETAAPBHGTYG 1329
DB 182 ECKKDDALELQSHAKSPSPVEREMVQVEREKTEAEPHVNVEKLEHETAVTVESEVSK 241
QY 1330 KPVLTLDMPSSERKALSLGGS--PSLPDQKACIEVQVOSLDTTYQTAAVKEVI-- 1386
DB 242 QLOTVNVPPIIDGAKEVSSLEGSPPCLGGEAVCTKIQVOSSESFTLTAABEEKVLG 301
QY 1387 ETVVISEGESPECVGAHLLPAEKSSATGHWTLQHAEDTVPLGPESQASIPITVPAP 1446
DB 302 ETANLLETGETLEPAGALVLEEKSEKNEDEPAHNGEDAVPTGPDCAKSTPVIYSATT 361
QY 1447 ESTLHPDLQGEISASQERSEEDK 1471
DB 362 KKGSSDLEGEKTTSLKWKSDDEVDE 386

```

```

RESULT 11
US-08-994-570-15
Sequence 15, Application US/08994570
Patent No. 6090929
GENERAL INFORMATION:
APPLICANT: Scott, John D.,
APPLICANT: Naurelt, Brian J.,

```

```

RESULT 12
US-08-973-462-8
Sequence 8, Application US/08973462B
Patent No. 6191270
GENERAL INFORMATION:

```

```

; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0860-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-8

Query Match          5.78; Score 462; DB 4; Length 1786;
Best Local Similarity 19.24; Pred. No. 5e-20;
Matches 325; Conservative 291; Mismatches 644; Indels 436; Gaps 69;

QY 53 QKNGQLSVNGVARGQDVHVOEENQEQE--BEVVDVGVQRESDVREK---DRVEE--- 105
DB 53 QVNGELAS-----BEVKEIILDLLEEGNTLTSESVDNNKNEEADIKENILLNIEBPKE 149

QY 106 -----MAANSTAVEDITKQGG-----EETSEIIEIQIPASENNVEENVQPAESQANDVG 153
DB 150 NIIDLNNIIONSEKQESVSNQVSDLPFNELLNSVDVNGEVKENILE--ESQVNDI 207

QY 154 PKKVFKFGFKPTVKKNKESQDTVQLLVKKDEGEGBASV--GAGDHQSPSVETAVGE 211
DB 208 FNSLVK-----SVQSQQHNVBEKVEESVEENDEESVEENVEENVDGSGVASSVEE 261

QY 212 S--ASKSELKQSTEQEGLKQE-----QSTEIPLQAESDQAAE 251
DB 262 SIASVDESIDSLSEENVAPTVVEIIVAPSVVSVESVAPSVESVEENVEENVEENVEE 321

QY 252 EAKDEGEKQKPTKSPSPSPVNSSETSSPKFFTHGWAGWRKKTSPKSKKEDDLE- 310
DB 322 NVEESVAENVEESVAENVEEIVAPTVVEIIVAPSVVSVESVAPSVESVEENVEE 381

QY 311 -TAKRKQEAKVDREEKTEPA---SEQEPAEDTDQARLSADYKVELPLDQVGD 366
DB 382 SVAENVEESVAENVEESVAENVEESVAENVEEIVAPTVVEIIVAPTVVEI 438

QY 367 LEASSEKCAPLATEVDEKMEAHQEVVAHVSTVEKEEQQGGEAGGVVEGTGE 426
DB 439 VAPSVSVAPSVESVEENVEE--ESVAENVEESVAENVEES-----VAENVEE 485

QY 427 SLPPEKLAEPQEVQEAPEALMKSRMCVSGGDHTQTLDSPEEKTLPKHPEGIVSEV 486
DB 486 SV-----AENVEESVAE-----NVEESVAENVEESVAENV 515

QY 487 EMLSSQBRIKVQGSLKPLFSSGLKGLKSKKQKRGKGQGGDEPBGYQHIIHTESPESAD 546
DB 516 EEIVAPTVVEIIVAPTVVEIIVAPSVVSVESVA-----PSVEESVEENVEESVA 560

QY 547 EQKGESSASSPEEP-----ETTCLKGLPEAPODGEABEGTTSDGKKRGGITPWASF 601
DB 561 ENVEESVAENVEESVAENVEESVAENVEEIVAPTVVEIIVAPTV-----BEIVAP--SV 612

QY 602 KMTVPKRRVRRPSSDKEEEL--EKVKSATLSSTDSITSEMDEVE--KTVGEEQKPEPKR 658
DB 613 ESVAPS--VEESVEENVEESVAENVEESVAENVEESVAENVEEIVAPTVVEIIVAP 665

QY 659 RVDTSVSWEALICVSSKKRARKASSDDEGGPRTLGGDSHRAEE--ASKDKAEAGTDVAP 716
DB 666 -----TVEEIVAPSVVSVESVAPSVESVEE-----NVEESVAENVEESVAENVE 708

QY 717 ASTQEQDQAQSSSPFAGSPSEGEVSTWESFKRLVTP-----RKKSCKLEEKAEDS 770

```

RESULT 13

US-07-853-913-2

; Sequence 2, Application US/07853913

; Patent No. 5338839

; GENERAL INFORMATION:

```

APPLICANT: McKay, Ronald D.G.
APPLICANT: Lendahl, Urban
TITLE OF INVENTION: Nestin Expression As An Indicator of
TITLE OF INVENTION: Neuroepithelial Tumors
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/853,913
FILING DATE: 19920319
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/660,412
FILING DATE: 22-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,803
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/201,762
FILING DATE: 02-JUN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/180,548
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-4641AAA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1805 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-853-913-2

Query Match 4.8%; Score 388; DB 1; Length 1805;
Best Local Similarity 21.3%; Pred No. 1.8e-15;
Matches 319; Conservative 233; Mismatches 543; Indels 402; Gaps 75;

QY 13 EOPAGSDT---PSELV-----LSHGPAAEASGAGDPAADADPATKLPOKNGQLSSVNGV 64
DB 440 EEPGQGGQGHFPDDTLATNLNPHHPTLEAKD--GESESRSVSIPODEQIWL--V 495
QY 65 AAGGDVHVOENEOG--EEVVDDEVGORESDEVAEKRVEMMAN---STAVEDITMD 119
DB 496 EKEADIEVVENSSAKTQESGLDTEQ--DSGGPLQKETLKLAGEPLMSLKIQYETA 554
QY 120 GGE-----ETSEIITBOIP---ASENNVEEMVQPAASSOANDGVFKKVFVFGF 163
DB 555 GKENCNSSTEGHGLTLEGPKEKQIPLKSLIEKNVE-----SEKTLENGVPVLSLLGK 608
QY 164 KETVKKDKNEKSPDTVQLTVKKDEGEAASVAGADHOEPSVET-----AVGESA 213
DB 609 EDTRTEDQMLSPKSG---TLKRSSLSGKESQEVVVRPSKGNLSWTAPEESQHPGLGPG 665
QY 214 SKESSELKQSTKEQGTU---KQESSTELPLQESDQ--AAEEADDEG-----EKGE 262
DB 666 AADQMLERLVEKEDQSPSPSEEDQACRPIQKENOEPLGYEEA--EGQILERLLEKES 723
QY 263 KEPTSPSEPPSSVNVSETTSSPKKFTTHGAGRKKTSPFKSKEDDLTFNAKKKEQEAEX 322
DB 263 KEPTSPSEPPSSVNVSETTSSPKKFTTHGAGRKKTSPFKSKEDDLTFNAKKKEQEAEX 322

```

```

DB 724 QESLRSPSEB---DQAGRSLOK-----ENQEPVGYEEADQMLERLLEKESQESLK 772
QY 323 VDEEKETEPASHEQ-----PAEDDQALSA-----DYEXVLEPLE-- 361
DB 773 SPENQIRIGPLERENKSLRYLEENQETVPLESRRQRLRLSEVEEORIVKPLEKV 832
QY 362 --DQVGLLEASSEKCAPLATEVDEKMEAHQEVAVHVSVEKTEEEOGGGGEAGV 419
DB 833 SQDSLGL---AEENVQPLR---YLEEDDCINKSLLE-----DKTKSLGSLDRNGDS 880
QY 420 VV-----EGTGESL--PPPK-----LAEPQEVQEAEPAEIIMKSRKCVSGGDHTOL 465
DB 881 IIPQSEETQVSLRPPREEDQRIYNHLEKESQEFSSSEEEQVW---ERSLSEGNHSL 937
QY 466 TDLSPF---EKLTPKHPGIVSEVEMLSQERIKVQGSPLKLFSSSGIKLSSGKKOK 521
DB 938 SSVEKEQWVESQLEKESQSGSKLED--EQEFTF---GPLEK--ENAESLSRLAGQDQ-- 989
QY 522 KRGGGDEEPGEYOHITHESPESADQKGSASSPPEPBTCTLEKGPLAPQD--GEAE 580
DB 990 -----EKGKLFQETQTLRAVNGMAVSP-----EKVDEPLPKPLGNDQ 1030
QY 581 EGTSDSEKKEKRGITPMASPKQWTPPK--RVRRPESDSEELKVKYS---ATLSST 633
DB 1031 EIASRLGKENGESLVSLEK--KGIEFTVKSLETEIIEPLET--AEEDLERKKSIDTQEPIMST 1088
QY 634 D---STYSEMDE---VKTYGEOKP---EPPKRVDTSVSWEALICVSSKKRARKA 682
DB 1089 EVARETYPPPEDEPPGSLGVDENRRTLSLEKESQELSLGKNNVETRVEDSQCLQVE 1148
QY 683 SSSDDEGGPRTL-----GGDSHRAEASKDEAGTDAVPAST----- 719
DB 1149 EGLQEOHQESLREVKQELPSSGNOQRMEDVBEKAVQGAFLATTVGTEDKALHLRG 1208
QY 720 --QEQDAQ-----GSSSPFAGSPSFG---EGVSTWESFKRLVTPRK 758
DB 1209 QGGEENAAABGELLQDIYGAWSLSGSEPKQRPVPAELDNLLEGGA-----LEVPAQ 1261
QY 759 SKSLKEKAD--SSVRLSTEIRPSRESVSIKKP1PGRKKRACQKQOQATVEDSGP 816
DB 1262 SMPVEYTERDEEDRAQAGBDSTLEVTGLGAARTGLEL-----EQGVGLIED--P 1307
QY 817 VEINDDPNVPAVPLSEYNVVEREKMAOGNTLPLQLGAVVYS--BELSTLVHTVSA 875
DB 1308 RHFAREKALPPS---LGEESYKAKIAGLEBPGKEPKENAGLBDGILELPTTSSALECO 1364
QY 876 VIDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPVVEVETKDIIAET----- 926
DB 1365 GHEEESMEGEEEBAS-----LETSDBSGSDAPQRPPEETEDEGAQALGTAPKLL 1418
QY 927 -----PVLQTLPEGKAHDMWTSE-----VPTSAVNAVATETSEALRTHEVTEA 972
DB 1419 EPCSPILPILT-----DAHLQPAEGIOEAGWQPEAGSSALRVEREPFGIAGEIPEG 1471
QY 973 -----SGAEETTDWVSAYSQUTDSPDTTEEATP---VOEVSGLDTEEBERQOALIQ 1023
DB 1472 LQDWEEGRBE-----SEADDLQETLPDSTPLGLYLRASAPKWDLAGOR-----LS 1518
QY 1024 AVADRYKKE--SOVPATQTVORTGSKALEKVEVEEDSEVLASEK-----EKDMPKGP 1075
DB 1519 PGGDAGKEKEDWGPVAPAAQ-----GLSGPREEEBEGHGSJLSEFEFDLGTASLPLPVP 1573
QY 1076 VOEAQAEHLAQ-----GSEFGQATPREL--EVEEVYADVHVAT----- 1112
DB 1574 KEV--ADHVQGVPPVLPACWDOGESDGFADBEESGEEGEEDDEGASGAGQWMSG 1631
QY 1113 -----COYIKLOQMEQAVAPESSETLTDSETNGST-----PLADSPDADQTODETTID 1161
DB 1632 ASGGCKVQODIAQRQDPVOESVGVSGLWDDGLRGAAANVPALAMVQSQSAPESSGEES-- 1689
QY 1162 SODSKATAVROSQVTEEEAATAQKEBPSTLPNNVPAOEHEGEEBGRDVLBTQOELTAA 1221
DB 1690 ---ESASLGEERGQVTDH--LDAPEVTVSWVGVGDADVDIGQSPNLD----- 1732

```

QY 1222 AVPLAKTEVG---QEGEVDWLDGEKVEQEVEF-----VHSGPN 1258
Db 1733 SEQVNGRMENLEAQEGV-VLDGD---EDQELLQGEVGAUKVPLVASPVHLGPS 1785

RESULT 14
US-09-134-001C-5080
; Sequence 5080, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GFC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5080
; LENGTH: 3696
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5080

Query Match 4.5%; Score 365.5; DB 4; Length 3696;
Best Local Similarity 18.2%; Pred. No. 1.2e-13;
Matches 324; Conservative 309; Mismatches 694; Indels 449; Gaps 75;

QY 46 DPATKLPQKQNL-----SSVNGVAEQGDVHVQ-----EENQGEVEEVDVQGESDV 97
Db 1462 DAITSLSKINAQVVEKESARNAIEQAKTQQTQFINNDNDATDEEKEVANNILVIATKQKSL 1521

QY 98 REKDRVEMAANSTAVEDITKOGQEETSEIEQIPASENNVVEVQPAESQANDVGFKV 157
Db 1522 ---DNINSLSNND-VENAKVAGINEANVLPAVAKSKAKKIDDKLQAOINOIQHOT 1577

QY 158 FKVGFKFTVKDKNEKSDTVOLLTVKKDEGEAEASVGAGDHQEPSVETA----- 208
Db 1578 -----ATTEKEAAIQLANQKSNBEARTAIQN-----EHNNGVAQAKSNGIHEIE 1622

QY 209 -VGSASKESELKOSTKEQGTLLKOBOSSTEIPLOAESDQAAEBE-----AKDEGE 258
Db 1623 LVMPDAHKKSQAKQSIDNK---VNEQSNT---INTTPDATDEEKQKALDKLAKDAGY 1675

QY 259 EKQEKPTKSPSPSPNSPSETTSFKFFTHGWAGWRKTSFK---KSKEDDLE----- 310
Db 1676 NKVDQAQTNOQVSDAKTEADITINIQ-----ANVAKFSAVELDSKFPDLKQINA 1728

QY 311 --TAEKKEQEA-----EKVDEEKEKTEPASEEOE-----PAEDTD- 345
Db 1729 TPNATEEKQDAIQLRNGKRDVEKNLINDRRDNEVEQHKNGIQLQELTHANPTRKSDA 1788

QY 346 ---QARLSADYKVELPLEDQVGDLEASSEK-----CAPLATEV 382
Db 1789 LQELQTFISQTELIN-----NNKDATNEEKDEAKRLLEISKNKTTININQATNNQV 1841

QY 383 FDEKMEAHQEVVAEVHVSTVEKTEBEGGGGEAEGGV-VVEGTGESLPPEKLAEPQEVQ 441
Db 1842 DNAKDNMGNEIATIIIPATTI-KTDAKTAIDKKAQEQVVTIINGNNDATDEEK-ABEARKLVE 1899

QY 442 EA--EPAEELMKS-REMCVSGDHTQITDL-----SPEEKTLPKHPEGIYSEVEMLSQE 493
Db 1900 KAKTEAKSNITNSTEREVNGAKTNGLEKINNIQPSQTQKTNAK-----QEINDKAOEQ 1953

QY 494 RIKVQGSPLKKLFSFSSGLKSLGKQKRGKGDDGDEPGEYQHI-----HTESPESADEQK 549
Db 1954 LIQINNTP-----DATEEKGQATNRVNAQLAQAINNNAHSTQEVNEK 1999

QY 550 GESSAS-----SPEPEETTCLEKGLSAPQDGE--ABEGTTSDEGKKREGITPWASPKKM 603

Db 2000 TNSIATIKSVQPNVIRKPTAINSLITQEANNQKTLIGNDGNATDDEKE-----AAQQL 2051
QY 604 VTPK--KRVRRPSESDEBELEKVKSATLSS---TSTVSEMDEKVTGGE---EQKPEEP 656
Db 2052 VTQKLEQIKIHESITQDNQVNVKAQAITAIKLINANAHRQODAINILTLAESKSDI 2111
QY 657 KRRVDTSVSWE--ALICVGSKKRARKKASSDDGGPRTLGGDSHRA--BEASKDKAEGT 712
Db 2112 RANQDATTEEKTAIOSIDDTLAQAR-----NNINGANTNALVDLENLEDGKQL 2160
QY 713 DAVPASTQEODQAGSSP---EPAGSPSGEGVST---WESFKRL----- 752
Db 2161 QRVLSTQTTQAKADIAQIGQORSTIDQONATTEEQEALERLNOETNGVNDRIQAA 2220
QY 753 -----VTPRKKS-----KSKLEEKAEDSSVVEQ---LSTEIPSPSEESVSI 790
Db 2221 LANQNVTDENKNNILETIRNVEPIVVKPANEIIRKAAEQTTLINQONQATLEEKQIAL 2280
QY 791 KFTIPGRRK-----KRADGKQEQATVEDSGPVEINEDDPNVAVPVPLSEYNAVEREKMEA 845
Db 2281 GKLEEVKNEALNQVSAHNSNDVKIAENNGIAKISEVHPET-----IIRKNA--KQIEQ 2333
QY 846 QGNTELPQLLGAVVVSSELSKTLVHTVSVAVIDGTRAV-----TSVEE 888
Db 2334 DAOSQIDTINANKSTNEEKSAADIRVNVAKIDAINNITNATTQLVNDKNSGNTSISQ 2393

QY 889 RSPSWISASVTEPLEHTAGEMPVEEVEKDIIAETPVLTLQTPGKDAHDMVTSEV 948
Db 2394 ILPS--TAVKTNALALASA-----KKNKAIIDOTPNATAEEKEANKVDRLOEEA 2444

QY 949 DFTSEAVTATETSEALRTEEVTEASGAE-----ETTDMVSAVSQJLTDSPDTTEATPVOEV 1004
Db 2445 DANILKAHTTDEVNNIKQAVQINAVQEVIKQNVQNLQNFIDNOKKIIENTPDATL 2504

QY 1005 ESGVLTDEEEROTAILQAVADKVKESQVPATQTVORTGSKALEKVEVEEEDSEVLAS 1064
Db 2505 E-----EKAEANRLQNLVLTSTSEI---ANVDHNEVDQALDKARPKIEBI---VPQV 2552

QY 1065 EKEKDVMPKPGVQAGAEHLAQSGEQATPESPESLEVPVADVHDVATCOVILKQQLMEQ 1124
Db 2553 SKRDVL--NAIQEAFNSQTEIQENQEAETNEB-----KTEALNKINQLNQ 2597

QY 1125 AVAPESSETLTDSETNGSTPLADSTADGTQDDDETIDSQDSKATAAVRQSQVTEE----- 1179
Db 2598 AKV-----NIDQASNKVDSDAKTSRSIQDIEQIQPHQPTKATG 2635

QY 1180 -----EAATAQKEPSTLPNNVPAQEBEGEPGRDVLPTQOELTAAAVPVLAKEVGOE 1234
Db 2636 RHLNEKANQQOSTIATHPNSTTEERQESAKLOEVLK-----KAIKIDKQGT 2684

QY 1235 GEVDWLDGK-----VKEOEVEFVHSGPNSQKAADVTDSEVMGVAGCQESKESTEVSLS 1289
Db 2685 ND-----DVEKTVVNGTAEIENILPATTVKDKAKADV-----AEKQKQLQINS 2729

QY 1290 LEEGEMETDVEKRETKPEQVSEEGEQETA-APHEGTYGKPVLTLDMPSSERKALGS 1348
Db 2730 NDEATTE---EKLVASDNLNHHVETTNOALEDAPD-----TNQNVNEKNKIGT 2775

QY 1349 LGG-SPSLPDQKA-GCIEVQVQSLOTTVTQTAEAV-----EKVIETVVI 1391
Db 2776 IRDIQLVVKKPTAKGIESAVEKKEKTEINQTNATHDEVREGLNQLNQHAKAKDNVQ 2835

QY 1392 SETGESPECVGAHLLPAEKSSGATGHWTLOHAEDTVPGLPESQAESIPITVPAPESTLH 1451
Db 2836 SQTNQVE-----NAEQNS-----LDQINNFRP--DFSKKRNAVAEIVKAQQNKI- 2878

QY 1452 PDLQGISASQRBRS-----EBE-----DKPAGPADGKESTAIKVLKAEPEIL 1497
Db 2879 DEIEQFSATQEEKDNALQHLDOVKIEINSINQANTDNEVDNAKTSGLNITEYREPY- 2937

QY 1498 ELESKKNKIVLNIQTAVDQFARTETAP-----ETHAYDSQTVQFACRLDSR----- 1544

Db 2938 ---NKKNAIILKLDVSTOEAINGYDPDATEDELOEANSKLINLLAKKQIGLAHTNN 2994

Qy 1545 EPNRCWTMKDKMKHPVPOPREDLQVLTLEAMAO 1580

Db 2995 EVDIYNEVSO-KMKITLPRVDTKAVARSVLNALAK 3029

RESULT 15

US-08-466-390-4

Sequence No. Application US/08466390

Patent No. 5686562

GENERAL INFORMATION:

APPLICANT: TOUKATLY, GARY

APPLICANT: LIDGARD, GRAHAM P

TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE

TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ & THIBEAULT

STREET: 125 HIGH STREET

CITY: BOSTON

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,390

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: PITCHER ESQ, EDMUND R

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: MTP-013

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 2101 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-466-390-4

Query Match 4.5%; Score 361.5; DB 1; Length 2101;

Best local Similarity 19.6%; Pred. No. 9.6e-14;

Matches 364; Conservative 292; Mismatches 668; Indels 529; Gaps 81;

Qy 53 QKNGGLSSVNGVABEGDVHVGENC--EGQEEVVDVGVGSESDVREKRV----- 103

Db 357 EKQAOLEKELSAALDDKCKLEKNEITLQGLSQ--LEBHLISQLQDNPPOKEGVLGDLVL 415

Qy 104 -----EEMANSTAVED-----ITKDGQSEET-----SEIIIOIPASE 135

Db 416 ETLKQEAATLAANNNTQLOARVEMLETERGQQAKLAEHGFHEEKEQJLSLITLQSSI 475

Qy 136 NNV---EEMVQPAESQA-----NDVGFKVYFKVFGKFTVKKDKAEK 174

Db 476 SNLSQAKELBEOQAHGARTLAQVASLTSELTLNATIQODDOELAGLQQAQKEKQAO 535

Qy 175 SPTVOLLTYKKBEGGAGSVAGDHQBPSTVAVGESAKSESELKOSTEKEGTLQKQ 234

Db 536 AQTLD-----QOEQASQGL-RHQ--VEQLSSSLKQKEQQLKEVAKEQEAAT--RQD 580

Qy 235 SSTEIPLQESDQAA--EEBAKDEGEKEKEPTKSP--SPSSPVNSETTSSFKKFTH 290

Db 561 HAQQLATAAEEEBASLRENDALQKLEAEKKAQKLTLQOOLQVANARASQAQTSVQ 640

Qy 291 GWAGRRKKTSPFKSKED--DLETAERKRKEQAEKVDEE---KEKTEPASEBOEPADT 344

Db 641 ---AQREKAEILSRKVELEQACVETARQEQHBAQNAVELEQLBSQCKATEKERVQEK 697

Qy 345 DQARLSADYKVELPLBDQVGDLEASSEBKCAPLATEVPEKMEAHQEVAAV-----H 398

Db 698 DQ--LQEQLOALKEKSLKVTGSLG---EEK--RRADALBEQQRCSIELNAETRLSLVEQH 750

Qy 399 VSTVEKTEBEEQCG-----GGEAEGGVVVGTEGSLPPEKLAEPQ----- 437

Db 751 KRERKELEEBERAGKGLBARLLQGEAHQAETELRLRELAEAMAAQHTASBECQLVKEV 810

Qy 438 -----EVPOEAP-----AEELMKSRMCVSGDHQTQDLSPEKTLPRKPEGIVS 484

Db 811 AAMDGYEDSQOEBAQYGANFOEQMLTKBECENARQLO---EAKEV-----AGIES 861

Qy 485 EVELMSQERIKYQSPKLKLF--SSGLKLSGKKQKGRGGGDEPBYQHITHESP 542

Db 862 HSELQIRQGNK-----LAEHLNHLARALQVQVEKVARQK--LADDLSTLQEKMAATSK 914

Qy 543 ESADEQKGESSASSPEEPTTCL--KGPLA---PDGEABEGTTSQGEKKRBGITP 596

Db 915 EVA---RLETLVRKAGEQOETASRELVEKPARAGDRQPEWLEEQG-----RQFCST 963

Qy 597 MASFKKWTTPKKRVRRPSESDEXE---ELEKVSATLSTSDTSYSE---MODEYKVTGEE 650

Db 964 QALQAM-----ERABQMGNELEKRLRALMESQGOQOQOEBERGQOEREARLITOE 1012

Qy 651 QKPEEPKRVDTSVSWBALICVSGSKRRARAKSSDDEGGPRTLGDSSHRAEASDKDE- 709

Db 1013 RGRQADLALEKARAELEBRLQNALNBOHVEFATLQELAHNL-----TEKSGXQEL 1066

Qy 710 ---AGTDAV-----PASTOEOQAQSSSPBPAGSPSBCGEVSTWESF 749

Db 1067 AKRLGLEAAQIKLELEBRLQTVKQLEQALAKKEKHAQSGSQAARTEPTGPKL-EAL 1125

Qy 750 KRLVTPRRKSKLEEKAB-----DSSVEQLSTEIRPSRESVSIKXF 793

Db 1126 RAEVSKLEBQCCQKQOQBQDLSLESLERLAEPRASRAERDNLLETLOGLQLEKQALCHQSQAL 1185

Qy 794 IPGRKKRKA--DGKQEOATVEDSGPVEI---NEDDPNVPAVPL--SEYNAVEREKMEAQ 846

Db 1186 ASAQRELAAPRTKYQDHSKADENKQAVARGQEAERKNSLISLEEBVSLANQVLEKE 1245

Qy 847 GNTLPQLGAVVYSELSKTLVHTVSVAVIDGTRAVTSVEERPSMTASVTEPLEHTA 906

Db 1246 G--ESKELKRLVMAESESQCLEBSCACC-----KQRPQA----- 1278

Qy 907 GEAMPVEEVTEDIIEETPVLTQTLPEGDADHDWYSE---VDFTEAVTATETSEA 963

Db 1279 -----TPELQNALMLGCR---RCPASGREAEKQKRVASENLRQELTSQAEKRAEBLQOE 1328

Qy 964 LRT---EEVTEASGAETTTDMVASVSLTDSPTTEATPVQEVESGVLDTEEBERQTOAI 1021

Db 1329 LKAMQEKFPQEQQLSTLQ-----EHTSTQALVSELLPAHGLQOQLQAE 1373

Qy 1022 LQAVADVKYKEEQVPATQTVORTGSKALE-----KYEEVEDSEVLASE 1065

Db 1374 QAAAEKRREBEL---OSKQAGALRAELRLAORELIGELPLRQKYAAEQERTAQQLRAE 1429

Qy 1066 K-----EKQVMPKGPQOEGAE 1082

Db 1430 KASYABQLSMLKKAHGLIAENRGLGERANIGROFLLEVELDOAREKVVQELAAARDAET 1489

Qy 1083 HLAQGSFTGQATPESLEVEVYTAADVAVATQVTKLOOLMEQAVAPESSETTLVTSSETGS 1142

Db 1490 RLAEVQRAQSTARELEV--MTAKYBGA-----KVKTLERQRFQERROKTLTNOVELS 1541

Qy 1143 TPLADSTADGTQOD--ETIDSQ--DSKATPAVROSQVTEEBAATAQKEBPSTLPNNVPA 1198

Db 1542 KKLADSDQASKYQOQKLAQVAGQSGEQEAFQAOQLNELQAOALSQEQAA----- 1593

Qy 1199 QEHGHEEFGRLVLEPTQOELTAAAVPVLAKTVEGQEGVDMLDEKXVKEBQEVFVHSGPN 1258


```

Db 1594 -----BHYKLQMEKAKTHYDAKKQONQELQELRSLEQLQKE-----N 1631
Qy 1259 SQAADVTYDSEVNGVAGCQKESTEV-QSLSLEEGEMETDV---EKEKRETKPEQVSEE 1314
Db 1632 KELRAEAEERLGHQLQAGLTKKEAQTGRHLTAQVRSLEAQVAHADQQLRDLGKFQVA-- 1689
Qy 1315 GEQETAAPHEGTYGKPVLTLDMP-----SSERKALG-----SLGGSPSLPD 1357
Db 1690 ----TDALKSREPOAKPQLDLSIDSLSCBEGTFLSITSKLPRTPQPDGTSVPGEPA SPI 1745
Qy 1358 QDKAGCIEVQVQSLD---TTVTQTABA-VEKVIETV--VISETGESPE----- 1399
Db 1746 SQR---LPPKVESLESYFTPIPARSQAPLESSLSDSLGDVFLDSGRKTRSRARRTTQIIN 1802
Qy 1400 -CVGAHLLPAFKSATGHHWTLQHAEDTVPLGPESQA-----ESIPILVTPAPESTL 1450
Db 1803 ITMTKLDVEEPDSANGSFYSTRSA-----PASQASLRATSTQSLARLGSPDYGN SA 1855
Qy 1451 HPDLQ--EISASQRESEEDKPDAGPDADGKESTAIEKVLAEP-----ILELES 1501
Db 1856 LLSLPGYRPTTRSSARRSQAGVSSGAPP---GRNSFYM-GTCODEPEQLDDWNR IAE LQ 1911
Qy 1502 KSNKIVLNVICTAVDQPARTETAPETHAYDSQTQVPACRLDSREPNCWTMMKDAKMKHP 1561
Db 1912 R-NRVCPHLKTCYP-----LESR-PSLSLGTITDEEMKTG 1945
Qy 1562 VPQP---REDLQVLTVLEAWA-----QPRK-----CLPR 1587
Db 1946 DPQETLRRASMQPIQIAEGTGITTRQQRKRVSLRPHQPGTPEKSKATSCFPFR 1998

```

Search completed: December 13, 2002, 00:17:10
 Job time : 54 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:16:06 ; Search time 224 Seconds
(without alignments)
118.719 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSSTEQRSPEQPAGSDT.....AWAQRKCLRLQLKAPVSK 1596

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105981 seqs, 16662342 residues

Total number of hits satisfying chosen parameters: 105981

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTU5_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8073	100.0	1596	9	US-09-902-432-4
2	5421.5	67.2	1346	9	US-09-902-432-2
3	4294	53.2	1781	10	US-09-738-877-3
4	462	5.7	1786	9	US-09-742-096-3
5	429	5.3	2665	10	US-09-864-761-34248
6	407.5	5.0	2368	10	US-09-815-242-5635
7	407.5	5.0	2368	10	US-09-815-242-12389
8	388	4.8	1325	10	US-09-864-761-35612
9	382	4.7	1879	9	US-09-971-536-70
10	377	4.7	2478	10	US-09-815-242-5816
11	377	4.7	2478	10	US-09-815-242-12967
12	377	4.7	3256	10	US-09-919-172-98
13	368	4.6	1400	10	US-09-764-176-7
14	354	4.4	1618	9	US-09-963-875-1
15	351.5	4.4	617	10	US-09-864-761-36182
16	339.5	4.2	2139	10	US-09-727-384-6
17	334.5	4.1	2344	10	US-09-815-242-12713
18	323.5	4.0	1884	10	US-09-785-770A-17
19	323.5	4.0	1907	10	US-09-785-770A-16

20	322	4.0	2310	9	US-09-991-496-120	Sequence 120, App
21	322	4.0	2310	10	US-09-874-923-120	Sequence 120, App
22	314	3.9	1863	9	US-09-734-672-4	Sequence 4, Appli
23	313.5	3.9	2843	9	US-09-987-482-1	Sequence 1, Appli
24	312.5	3.9	2843	8	US-08-681-219-32	Sequence 32, Appli
25	308.5	3.8	26926	9	US-09-759-508B-2	Sequence 2, Appli
26	306	3.8	1863	9	US-09-734-672-2	Sequence 2, Appli
27	306	3.8	1863	9	US-09-734-672-6	Sequence 6, Appli
28	305	3.8	3158	10	US-09-815-242-12611	Sequence 12611, A
29	304	3.8	6281	10	US-09-815-242-12996	Sequence 12996, A
30	296.5	3.7	1285	10	US-09-982-091A-2	Sequence 58, Appli
31	292.5	3.6	350	9	US-10-101-487-58	Sequence 58, Appli
32	290.5	3.6	1597	9	US-09-832-292-35	Sequence 35, Appli
33	285.5	3.5	2437	10	US-09-815-242-5834	Sequence 5834, Ap
34	284	3.5	1501	10	US-09-924-154-17	Sequence 17, Appli
35	281.5	3.5	1065	10	US-09-864-761-43248	Sequence 43248, A
36	279.5	3.5	2828	10	US-09-905-129-21	Sequence 21, Appli
37	279.5	3.5	2828	10	US-09-991-630-21	Sequence 21, Appli
38	278	3.4	2076	10	US-09-815-242-5815	Sequence 5815, Ap
39	278	3.4	2186	10	US-09-815-242-12913	Sequence 12913, A
40	277.5	3.4	2025	10	US-09-815-242-5703	Sequence 5703, Ap
41	274	3.4	868	9	US-09-884-001-19	Sequence 19, Appli
42	267.5	3.3	1332	10	US-09-982-091A-4	Sequence 4, Appli
43	266.5	3.3	1192	10	US-09-758-140-6	Sequence 6, Appli
44	266.5	3.3	1192	10	US-09-972-599A-6	Sequence 6, Appli
45	265.5	3.3	1192	10	US-09-789-386-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-902-432-4
; Sequence 4, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1596
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-902-432-4

Query Match 100.0%; Score 8073; DB 9; Length 1596;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGAGSSTEQRSPEQPAGSDTPELVLSGHGPAAEASGAAGDPADADPATKLPQKNGQLSS	60
DB	1	MGAGSSTEQRSPEQPAGSDTPELVLSGHGPAAEASGAAGDPADADPATKLPQKNGQLSS	60
QY	61	VNGVAGQGVHVHVEENQEGEEVVDVQRESEDEVREKORVEEMAA NSTAVEDITKDG	120
DB	61	VNGVAGQGVHVHVEENQEGEEVVDVQRESEDEVREKORVEEMAA NSTAVEDITKDG	120
QY	121	QETSETIEIOIPASENNVEMVQPAESQANDVGFKVFGFKFTVKDKNEKSDTVQL	180
DB	121	QETSETIEIOIPASENNVEMVQPAESQANDVGFKVFGFKFTVKDKNEKSDTVQL	180
QY	181	LTVKDEGEAEASVGAGDHQEPSEVETAVGESASKESELKQSTKEGTLKQESSTEIP	240

|||||
Db 181 LTVKKDBEGEADASVAGDHOEPPVETJAVGASASKESELKOSTEQEBTLQOEGSTIIP 240
Qy 241 LQASDQAABEABADEGEKEKEPTKSPESPSPVNSSETTSFFKFFTHGAWGRKXTS 300
Db 241 LQASDQAABEABADEGEKEKEPTKSPESPSPVNSSETTSFFKFFTHGAWGRKXTS 300
Qy 301 FKSKEDDLETAERKKEBAEKVDEEKEKTEPASEBOBPADTQALSLADYKVEIPL 360
Db 301 FKSKEDDLETAERKKEBAEKVDEEKEKTEPASEBOBPADTQALSLADYKVEIPL 360
Qy 361 EDVQDLEASSSEKCAPLATEVFEDEKMEAHQEVVAEYVSTYKTEEBOGGGGEAGGV 420
Db 361 EDVQDLEASSSEKCAPLATEVFEDEKMEAHQEVVAEYVSTYKTEEBOGGGGEAGGV 420
Qy 421 VEGTGESLPEPEKLAPOEVPOEABPAEELMKSRMVCSGDHTQTLTDLSPBEKTLPKHPE 480
Db 421 VEGTGESLPEPEKLAPOEVPOEABPAEELMKSRMVCSGDHTQTLTDLSPBEKTLPKHPE 480
Qy 481 GIVSEVEMLSOERIKVQGSPLKLPSSSGLKLSGKKQKGRGGGDBEPGEYOHITE 540
Db 481 GIVSEVEMLSOERIKVQGSPLKLPSSSGLKLSGKKQKGRGGGDBEPGEYOHITE 540
Qy 541 SPESADBOGESSASSPEPEETTCLEKGPLAPQDGEABEGTSDGEKRGGITPMASF 600
Db 541 SPESADBOGESSASSPEPEETTCLEKGPLAPQDGEABEGTSDGEKRGGITPMASF 600
Qy 601 KKMVTPKRRVRPSESDKEELEKYSATLSSTDSVSEMDEVTVGEBOKEPBEKRV 660
Db 601 KKMVTPKRRVRPSESDKEELEKYSATLSSTDSVSEMDEVTVGEBOKEPBEKRV 660
Qy 661 DTSVSWMALICVSSKKCARAKASSDDDEGPTTLGGDSHRAEASKXKAGTDAVPASTQ 720
Db 661 DTSVSWMALICVSSKKCARAKASSDDDEGPTTLGGDSHRAEASKXKAGTDAVPASTQ 720
Qy 721 EDDQOQSSSPBPASSPEBEGEVSTWESFKRLVTPRKSKSKLBEAKEDSSVEQLSTEIE 780
Db 721 EDDQOQSSSPBPASSPEBEGEVSTWESFKRLVTPRKSKSKLBEAKEDSSVEQLSTEIE 780
Qy 781 PSRESWWSIKKFIPIGRKRKADGQEOATVEDSGPVEINEDPNVPAVPLSEYNAVER 840
Db 781 PSRESWWSIKKFIPIGRKRKADGQEOATVEDSGPVEINEDPNVPAVPLSEYNAVER 840
Qy 841 EKMEAQNTLEPOLLIGAVVSELSKTLVHTVSVAVIDSTRAVTSYEERSPSMISASYTE 900
Db 841 EKMEAQNTLEPOLLIGAVVSELSKTLVHTVSVAVIDSTRAVTSYEERSPSMISASYTE 900
Qy 901 PLEHTAGEAMPPEVEVTEKDIIAETPVLTOQLPBGKAHDMVTSEVDTSEAVTATET 960
Db 901 PLEHTAGEAMPPEVEVTEKDIIAETPVLTOQLPBGKAHDMVTSEVDTSEAVTATET 960
Qy 961 SPALTEETVEASGABETTDWMSAVSOLTDSPDTEATPVOEBSGVLDTEEBERQTA 1020
Db 961 SPALTEETVEASGABETTDWMSAVSOLTDSPDTEATPVOEBSGVLDTEEBERQTA 1020
Qy 1021 ILQAVADKVKEESQVPATQVORTGSKALEKVEVEBDESEVLASEKXOVMFKGPYQENG 1080
Db 1021 ILQAVADKVKEESQVPATQVORTGSKALEKVEVEBDESEVLASEKXOVMFKGPYQENG 1080
Qy 1081 AEHLAQSGSTGATPESLEPEBVTADVHVATCOVTKLOQLEMAQVAPSSSETLTDSEIN 1140
Db 1081 AEHLAQSGSTGATPESLEPEBVTADVHVATCOVTKLOQLEMAQVAPSSSETLTDSEIN 1140
Qy 1141 GSTPLADSTAGTOODETIDISQDSKATAVAVOSQVTEBEAATAQKEBETSTLPNNVPAOE 1200
Db 1141 GSTPLADSTAGTOODETIDISQDSKATAVAVOSQVTEBEAATAQKEBETSTLPNNVPAOE 1200
Qy 1201 EHGEPEGRDVLPTQOELTAAAVPLAKTEVQEGEVMDLGEKVEBOEVVHSGPNSQ 1260
Db 1201 EHGEPEGRDVLPTQOELTAAAVPLAKTEVQEGEVMDLGEKVEBOEVVHSGPNSQ 1260
Qy 1261 KAADVTYDSEWVGAVAGCQKEKSTEVQSLSEBEGMETDVEKREKTEPQEVSEBQETA 1320
|||||

Db 1261 KAADVTYDSEWVGAVAGCQKEKSTEVQSLSEBEGMETDVEKREKTEPQEVSEBQETA 1320
Qy 1321 APEHEGTGKPVLTLDMPSSERKALSLGSPSLPPODKAGCTIEVYQSLDITVTQTAE 1380
Db 1321 APEHEGTGKPVLTLDMPSSERKALSLGSPSLPPODKAGCTIEVYQSLDITVTQTAE 1380
Qy 1381 AVEKVIETVWISSETGESPECEYGAHLPAEKSSATGHWTLQHAEDTVPLGPESQAESIPI 1440
Db 1381 AVEKVIETVWISSETGESPECEYGAHLPAEKSSATGHWTLQHAEDTVPLGPESQAESIPI 1440
Qy 1441 IVPAPSTLHPDLQGEISASQERSESEBDKPADPADGKESTAIKVLKABPELLELE 1500
Db 1441 IVPAPSTLHPDLQGEISASQERSESEBDKPADPADGKESTAIKVLKABPELLELE 1500
Qy 1501 SKSNKIYLVNIQTAVDOPARTETAPETHAYDSQTVPAACRLDSNEPRCWTMADAKMKA 1560
Db 1501 SKSNKIYLVNIQTAVDOPARTETAPETHAYDSQTVPAACRLDSNEPRCWTMADAKMKA 1560
Qy 1561 PVPQREDLOVLTVLEAMAOPRKCLPRLQAKAPVSK 1596
Db 1561 PVPQREDLOVLTVLEAMAOPRKCLPRLQAKAPVSK 1596

RESULT 2
US-09-902-432-2
; Sequence 2, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Susan G. Jaken
; APPLICANT: Irwin H. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A0558-A-FWC-A 070156.0597
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1346
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-902-432-2

Query Match 67.2%; Score 5421.5; DB 9; Length 1346;
Best Local Similarity 91.6%; Pred. No. 6.4e-229;
Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;

Qy 387 MEAHQEVAAEHVSTVEKTEEBEBOGGGGEAGGVVEGTGSLPEPEKLAPOEVQEAEP 446
Db 1 MEAHQEVAAEHVSTVEKTEEBEBOGGGGEAGGVVEGTGSLPEPEKLAPOEVQEAEP 446
Qy 447 EELMKSRMVCSGDHTQTLTDLSPBEKTLPKHPRGIVSEVEMLSOERIKVQGSPLKLP 506
Db 61 EELMKSRMVCSGDHTQTLTDLSPBEKTLPKHPRGIVSEVEMLSOERIKVQGSPLKLP 506
Qy 507 SSSGLKLSGKKQKGRGGGDEBEGEYOHITHSPESADQKESASSPEPEETTC 566
Db 121 SSSGLKLSGKKQKGRGGGDEBEGEYOHITHSPESADQKESASSPEPEETTC 566
Qy 567 EKGPLAPQDEAELEGTTSDGEKRE---GITPMASFKKMVTPKKRVRRPSESDEBEL 622
Db 181 EKGPLAPQDEAELEGTTSDGEKRE---GITPMASFKKMVTPKKRVRRPSESDEBEL 622
Qy 623 EKVSATLSTSDTSVSEMDEVKTVEGEBOKEPBEKRVDTSVSMALICVSSSKKARAK 682
Db 225 EKVSATLSTSDTSVSEMDEVKTVEGEBOKEPBEKRVDTSVSMALICVSSSKKARAK 682
Qy 683 SSSDDEGPTTLGGDSHRAEASKXKAGTDAVPASTQEOQAGSSSPBPAGSFBEG 742
|||||

```
Db 295 SSSDIR-GPRTLGGGQSRGQQRSDTAVPASTQEQQAQSSSPAGSPSGEG 353
Qy 743 VSTWSEFKRLVTPRKKSLEEK-ABDSSVEQLSTBIEPSREBSWS--IKKPIPQRK 799
Db 354 VSTWSEFKRLVTPRKKSLEEKAGRTLVVAGGCPRLSRVKEKNLGFPLRNSSPDGR 413
Qy 800 KRADGQEOATVEDSGVEINEDDPNVPVAVPLSEYNAREKMEAGNTLPLLGA 859
Db 414 KQMGROQATVEDSGVEINEDDPNVPVAVPLSEYNAREKMEAGNTLPLLGA 472
Qy 860 VSELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPPVEE 919
Db 473 VSELSKTLVHTVSVAVIDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPPVEE 532
Qy 920 DIIAETPVLVTLPEKGADHDMVTSEVDFTSSEAVTATSETSEALRTEEVTEASGAE 979
Db 533 DIIAETPVLVTLPEKGADHDMVTSEVDFTSSEAVTATSETSEALRTEEVTEASGAE 592
Qy 980 DMVSAYSQLTSDPTTEATPVQVEVSGVLTDEEERQTAILOAVADKVKESQVPATQ 1039
Db 593 DMVSAYSQLTSDPTTEATPVQVEVSGVLTDEEERQTAILOAVADKVKESQVPATQ 652
Qy 1040 TVORTGSKALEKVEEVEDSEVLASEKEDVMPKGPVQEAAGHLAGSETGOATPRESLE 1099
Db 712 TVORTGSKALEKVEEVEDSEVLASEKEDVMPKGPVQEAAGHLAGSETGOATPRESLE 712
Qy 1100 VPEVTADVHVATCQVTKLOLMEQAVAPESSETLTLDSETNGSTPLADSDTADTQDET 1159
Db 713 VPEVTADVHVATCQVTKLOLMEQAVAPESSETLTLDSETNGSTPLADSDTADTQDET 772
Qy 1160 IDSQDSKATAAARQSVTESEAAATAQKEEPTSLPNNVPAQEEHGEERGRDVLPTQO 1219
Db 773 IDSQDSKATAAARQSVTESEAAATAQKEEPTSLPNNVPAQEEHGEERGRDVLPTQO 832
Qy 1220 AAAPVPLAKTEVGOEGEVDMLDGEKVEEQEVFVHSGPNQKAADVTYDSEVMVAGCO 1279
Db 833 AAAPVPMQKTEVGOEGEVDMLDGEKVEEQEVFVHSGPNQKAADVTYDSEVMVAGCO 892
Qy 1280 KESTEVQSLSEEGEMETDVEKRETKPEQVSEGEQETAAPHEGTYGKPVLTLDMP 1339
Db 893 KESTEVQSLSEEGEMETDVEKRETKPEQVSEGEQETAAPHEGTYGKPVLTLDMP 952
Qy 1340 SERGKALGSLGGSPSLPDQKAGCIEVQVQSLDVTVTQTAEAVEKVITVVISETGSPE 1399
Db 953 SERGKALGSLGGSPSLPDQKAGCIEVQVQSLDVTVTQTAEAVEKVITVVISETGSPE 1012
Qy 1400 CVGAHLPLPAEKSSATGCHWTLOHAEDTVPLGPESQAESIPILVTPAPESTLHPDLO 1459
Db 1013 CVGAHLPLPAEKSSATGCHWTLOHAEDTVPLGPESQAESIPILVTPAPESTLHPDLO 1072
Qy 1460 ASQRESEEDKPDAGPDAGKSTAJEKLKAPPEILELESKNKVLNVIOTAVDQFA 1519
Db 1073 ASQRESEEDKPDAGPDAGKSTAJEKLKAPPEILELESKNKVLNVIOTAVDQFA 1132
Qy 1520 RTETAPTHAYDSQTPACRLDSREPNCWTMKMDAMKHPVQPPREDIQLVTLVLEAW 1578
Db 1133 RTETAPTHAYDSQTPAMRLDSREPNCWTMKMKAKMHPVQPPREDIQLVTLVLEAW 1191

RESULT 3
US-09-738-877-3
; Sequence 3, Application US/09738877
; Patent No. US20020015970A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan
; APPLICANT: Weiss, Stephen J.
; APPLICANT: Glynn, Richard
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF ANGIOGENESIS, COMPOSITIONS, AND MET
; TITLE OF INVENTION: SCREENING FOR ANGIOGENESIS MODULATORS
; FILE REFERENCE: A-69806/DJB/JJD
; CURRENT APPLICATION NUMBER: US/09/738,877
```

```
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT/US 00/22061
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-738-877-3

Query Match 53.2%; Score 4294; DB 10; Length 1781;
Best Local Similarity 57.9%; Pred. No. 1e-179;
Matches 959; Conservative 187; Mismatches 401; Indels 108; Gaps 35;

Qy 1 MGAGSSTEQRSPEQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 56
Db 1 MGAGSSTEQRSPEQ-PAGSDTPSELVLSHGHPAAEAS-GAAGDPADA--DPATKLPQKNG 60
Qy 57 QLSNVNGVAEQGDVHVQENQEQ-----EEVVDEVDQGESEEDREKDR 102
Db 61 QLSNVNGVAEQGDVHVQENQEQ-----EEVVDEVDQGESEEDREKDR 120
Qy 103 VEMAAANSTAVEDITKDGQEBTSEITQIPASENNVVEVMQPAESQANDVGFKKVFKFVG 162
Db 121 DKEMATKSAVVDITDDGQENRN-IEQIPSSSENLEELTQPTESQANDIGFKKVFVKFVG 179
Qy 163 FKFTVKDKNEKSDTVQLLTVKKDEGEAEASVAGDGHQEPSVETAVGESASKESELKQS 222
Db 180 FKFTVKDKKTEKPDVTQLLTVKKDEGEA--AGAGDHQDPSL--GAGEAASKESEPKQS 234
Qy 223 TEKQEGTLKQEQSSTIPIQAESDQAAEBEAKDEGEKEKEPTKSPESPSPNSETTS 282
Db 235 TEKPEETLKREQSHABISPPAESQAV-BECKEGBEKEKEPSKGAESPTSPVTSETGS 293
Qy 283 SPKKFTTHGWAGWRKTSFKKSKEDDLETAERKEQEAQKVDDEEKEKTEPASEE---- 337
Db 294 TFKKFTQGWAGWRKTSFRKPKEDVEASEKKKEQEBEPEKVDTEEDGKAIVASEKLTASE 353
Qy 338 ----QBPADTDQARLSADYKVELPLEDQVGDLEASSEEKCAPLATEVDFDERMEAHQ-E 392
Db 354 QAHQPQPAESAHEPRLSAEYKVELPSEEQVSGSQSQSEKPAPLATEVDFDEKIEVHQEE 413
Qy 393 VVAEVHVSTVEKTEERQGGGGEAGGVVVEGTGESLPPEKLABPQVPOPAEABELMKS 452
Db 414 VVAEVHVSTVEERTEQ-----KTEVEBETAGSVPAEELVGMDAEPQEAEPKLVKL 465
Qy 453 REMCVSGGDHDTQLDLSPEEKLTPKHPEGIVSEVEMLSQERIKVQGSPLKLFSSSGLK 512
Db 466 KETCVSGEDPTQQAADLSPEKVLKSPGEGVSEVEMLSQERIKVQGSPLKLFSTGLK 525
Qy 513 KLSGKKQKRGKGGGDEEPGEYQIHITESPESADEQKGESSASSPEPEBTTCLEKGPLE 572
Db 526 KLSGKKQKGRK-GGGDEESGEHTQVPADSPDQEEQKGESSASSPEPEITCLEKGLAE 584
Qy 573 APDQGEAEAGTTSDGKKKREGITPWAFFKMMVTPKRVRRRPSSEDKKEELEKVKSATLSS 632
Db 585 VQDQGEAEAGTSDGKKKREGITPWAFFKMMVTPKRVRRRPSSEDKKEELEKVKSATLSS 644
Qy 633 TDSTVSEMDEVKTVGEEOKEPEKPERVDTSVSWEALICVGSKKKARKASSSSDDGGPR 692
Db 645 TESTASEMQEEMKGSVEEPKPKPKVDTSVSWEALICVGSKKKARRRRSSDEGGPK 704
Qy 693 TLGGDSHRAEASKDXEAGTDAVPASTQDQQAQSSSPAGSPSGEGVSTWESFKRL 752
Db 705 AMGDQKADKADKXETGDTGLAGSQEHDPQCGSSSPQAGSPTEGEGVSTWESFKRL 764
Qy 753 VTPRKKSCKLBEKAEAD----SSVEQLSTEIPSRRESVSIKKFTPGRRKKRAGDQKQEQ 808
Db 753 VTPRKKSCKLBEKAEAD----SSVEQLSTEIPSRRESVSIKKFTPGRRKKRAGDQKQEQ 808
```

```

Db 765 VTPRRKSKLEEKSESDSIAAGSVHSTPDPBPKBESWVSIKKFIPGRKKRPPDQKQ 824
Qy 809 ATVEDSGVEINEDDPNNPVPVPLSEYNAAVEBKME--AQGNTLIPQLGAVYSEELS 865
Db 825 APVEDAGPFGANEDSDVPAVVPVLPSEYDAVEREKMEAOQAQGAQEPQEKATEVSKELS 884
Qy 866 KTLVHTVSAVVDGTRAVTVSEERSPSWISASVTEPLEBTAGEAMPVPEVEKIDIA-E 924
Db 885 ESQVMMMAAAVADGTRAAITIEERSPSWISASVTEPLEBVEKMAALLTEVLEVERITAE 944
Qy 925 ETPVLTQTLPECKDAHDDVNTSEVDFTSEAVTATETSEALRTEEVTAAGAEITDMVA 984
Db 945 BEPTVTEPLPENREARAGDVVSEAEITPAETAAETAGPLGSEBEGTAAAEETEMVA 1004
Qy 965 VSQLTDSPTTEATIPVQVESGVLDTEEBERQTAIIOAVADKYEESQVAT----- 1038
Db 1005 VSQLTDSPTTEATIPVQVEGVDPIDEBQRRTOEVLQAVAEKVEESQLTGTGPEDV 1064
Qy 1039 -QTVQRTSGKALEKVEEVEDESEVLASEKEKDVMPKGPVQEAAGAEHLAQSGETQATPS 1097
Db 1065 LQPVRA-----EAKRPEQAESGLKKEITVVLKVAQAEKTEPFTQGVVGGTTTPE 1118
Qy 1098 LE-VEEVTAADVH--VATCOV-----IKLOOL-MEOAVAPESSELTLDSETNGSTPLAD 1147
Db 1119 EKARQVTSISSSELYTTQCAETLAGVKSQEMMEQALPPDSVETPTDSETDGTSPVAD 1178
Qy 1148 SDTADGTQODETIDSDSKATAAVRQSVQTEBEAATAQKEBPSTLPNNVPAOEHEGEPG 1207
Db 1179 FAPGTQODEIVEIHEENEVASGTQSGTEAFAVPAQKERRPA--PSSFVQEBETEQSK 1237
Qy 1208 -RDVLEPQOELTAAAVPLATTEVQGEVDMQGEKXK-----BEQVFVHSG--PNS 1259
Db 1238 MEDTLEHTDKSVSVETVSIKSTEGTQ--EADQVADKTKDVPFFEGLSIDTGITVSR 1295
Qy 1260 OKAADVTYDSEVMGVAGCOEKSESTEVOS-----LSLEBEMETDVEKEKRETKPEQVSE 1314
Db 1296 EKYTEVALKGEETEBECKKDALBELQSHAKSPSPVEREMVQVREKTEAPTEPTVNEE 1355
Qy 1335 G-EQETAAEBHEGTYGKPVLTLDMPSSERKALIGSGS--PSLPDQKACIEVOYQSID 1372
Db 1356 KLEHETAAVTVSEVSKQLQTVNVPFIIDAKESVLESGPPQLQGEBAVCTKIQVQSSB 1415
Qy 1373 TTVTQAAEAVKVI--ETVVISGTSPECVGAHLLPAKSSATGSHMQLQHAEDVPIG 1430
Db 1436 AFTTLTAAEBEKKVIGETVNILETGTLPEAGAHVLEKSSSKNEDFAHPGEEDAVPTG 1475
Qy 1431 PPSQASIRPIYTPAPESLHPDLOGEISASQERSEEDKPDAGPDADGKESTALEKL 1490
Db 1476 PCCQAKSTVIYATTKKLSDBLEKETTSLKMKSDVDEVOVACQEV--KYSVAIEDL- 1532
Qy 1491 KAEPE--ILESSKSNKIYLVNIQTAVDQFART-ETAPETHAYDQTVQVPAICRLDREPN 1547
Db 1533 --EPENGILELETKSKLVONIQTAVDQFVRTEETATMLTSELQTAHVIAKADQDAG 1590
Qy 1548 KCMWK-----MDAKMKHPVQPREDLQVLTLEA 1577
Db 1591 QETKEGEPPOASADETPITSKESSESTAVGOA 1625

```

```

RESULT 4
US-09-742-096-3
; Sequence 3, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DRUIHE, PIERRE
; APPLICANT: DAUBERGES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894

```

```

; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-3

Query Match 5.7%; Score 462; DB 9; Length 1786;
Best Local Similarity 19.2%; Pred. No. 5.6e-13;
Matches 325; Conservative 291; Mismatches 644; Indels 436; Gaps 69;

Qy 53 QNQGQSVNGVAGQGVHVOENQGOE-EEVDEVDGQRESDVRK--DRVEE-- 105
Db 95 QVNGELAS-----EEVKEKIIDLLEEGNTLTESVDDNKNLEAEDIKENILSNIEPKE 149
Qy 106 -----MAANGTAVEDITKDGQ---EETSEIIEQIPASENNVEENVQPAEQANDVG 153
Db 150 NIIDLNNIGNQSEKQSVSENVQVSDLEFNEILNSVDVNGEYKENTLE--ESQVNDI 207
Qy 154 FKVPKFGFKFTYKDKNEKSDTVQLLTVKKDEGGAASV--GAGDHPSPVETAYGE 211
Db 208 FNSLYK-----SVQEQOHVVEKEVESVEENDESVVEENVEENDEDSVASSVEE 261
Qy 212 S--ASKESSELKQSTEKQGTIKQKQ-----QSTETPLQAESDQAAE 251
Db 262 SIASVDESIDSIEENVAFTVEIIVAPSVESVAPSVESVEENVEESVAENVEESVAE 321
Qy 252 EAKDGEKEQKEPFTKSPSSPNSSTTSFKKFTFHGAQRKKTSPKSKSEDDLE- 310
Db 322 NVESVVAENVEESVAENVEEIVAPFTVEIIVAPSVESVAPSVESVEENVEE 381
Qy 311 -TAERKQEAQKQDEEKEKTEPA---SEQEPADDTQALSDYKVELPLEDQVGD 366
Db 362 SVAEVNEESVAENVEESVAENVEESVAENVEESVAENVEEL-VAPFTVEIIVAPFTVEE- 438
Qy 367 LEASSEKCAPLATEVFDEKMAHQEVAEVAHVSTVEKTEEBQGGGABAGGVVEGTGE 426
Db 439 VAPSVESVAPSVESVEENVE--ESVAENVEESVAENVEES-----VAENVEE 485
Qy 427 SLPPKLAEPQEVPEAPAEELMKSRMCVSGGDHTQLTLPSPBEKTLPHHPBEIVSEV 486
Db 486 SV-----AEVNEESVAE-----NEESVAENVEESVAENV 515
Qy 487 EMLSQERIKVQGSPLKLFSSGKLKLGKKQKGGGGDEBEGEYQHITHSPESAD 546
Db 516 BEIIVAPFTVEIIVAPFTVEIIVAPSVESVA-----PSVESVEENVEESVA 560
Qy 547 EQKGSASSPREP-----EETTLCEKGPLAPQDGAEBGTTSDGKKRKGITPMASF 601
Db 561 ENVEESVAENVEESVAENVEESVAENVEEIVAPFTVEIIVAPFTV-----EIVAP- 612
Qy 602 KMVTKKVRKRPSESDKEEL-EKYKSAITLSSTDSVTSEMODEV--KTVGEOKEKEEKR 658
Db 613 BSVAPS--VEESVEENVEESVAENVEESVAENVEESVAENVEEIVAPFTVEIIVAP- 665
Qy 659 RYDTSVSMELICVSGSKRRKAKSSSDDEGGPRLGDSHRAEE--ASKQKACTADVP 716
Db 666 -----TVEIIVAPSVESVAPSVESVEE-----NVEESVAENVEESVAENVE 708
Qy 717 ASTQEQDQAGSSSPPEPASPSEBEGVSTWESFKLVT-----RKKSLSKLEKADS 770
Db 709 BSAVAVNEESVAENVEEIVAPFTVEIIVAPFTVEIIVAPSVESVAPSVESVEENVEES 766
Qy 771 SVEQLSTEIPEPSESWVSIKKFIPGRKKKADGQKQDA--TVEDSGVEINEDDPNPA 828
Db 767 VAENVEESVAENVEES-----VAENVEESVAPFTVEIIVAPSVEE-----S 806
Qy 829 VVPLSEYNAAVEREKMEAOQNTLIPQLGAVVYSEELSTLVHTVSAVVDGTRAVTSVEE 888

```

Db 807 VAPSVESVAENVATNLSNDL-LSNLLGGI-ETEEIKDSILNEIE-----EVKE 853
QY 889 RSPSWISASVTEPLEHTAGAMPVEVTEKDIAETPVLVTLPLPG-KDAHDDVMWTS 947
Db 854 N-----VVTILENEETTAESVTTFSNILEE---IQENTITNDTIBEKLEHENVLSAA 906
QY 948 VDFTEAVTATSEALRTEVEASCAEETDMVSAVSO-----LTDSPDTTEAT 999
Db 907 LENTQSEEEKKEVDVJ--BEVKE-----EVATTLIETVEQAESKANTITFEINDEE-- 958
QY 1000 PVOEVESGLVDTBEERQTOAILQAVADKVKERSQVPATQTVORTGSKA-----LEKVEE 1054
Db 959 --NAVESNENVAENLEKLNENFVTLVKVEETVEISGESLENEMDKAFSEIFDNVKG 1016
QY 1055 VEED-----SVLASEKEKOV-----MPKG-----PVQ 1077
Db 1017 IQENLLTGMFRSITSVIOSEKVDLNENVSVSILDNIENMKELKLNENISSTSEGVQ 1076
QY 1078 EAGAHLAQG-----SETQATPESLEVPEVTADVDHVATCOVTK-- 1117
Db 1077 ETVEHVEQNVYDVDPAMKDOFLGILNEAGGLKEMFFNLEDVFKSESVDITVEEIKDE 1136
QY 1118 -LOQLMEQ--AVAPESSETLTSETNGSTPLADSDTADGTQODETIDSDQSKATAVRQ 1173
Db 1137 PVQKEVEKETVSIIEEMEENVDVLEEKEDLTK-MIDAVEESIEISSDSKEETESIKD 1195
QY 1174 SQYTBEEAATAQKEPSTLNNVPAQEHGEPEGRVLE--PTQOELTAAAVPV----- 1225
Db 1196 K--EKDVSJLVBE-----VQNDMDSEVKEVLELKNMBEELMKDAVEINDITSK 1242
QY 1226 -LAKTEVQGEVDWL--DGEKVKE-----EQEVFVHSGPNSQKAADTYD-----SE 1270
Db 1243 LIBETOELNEVEADLIKDMEKJLEKALSSEDSKEIIDAAXDDTLKVEIEEHDTITLDE 1302
QY 1271 VMGVAGCQE-----KESTEVQSLSEEGE---METDVEKEKRET 1306
Db 1303 VVELKDVDEDKIEKVSLDKLEEDILKEVKEIKELESEILEYKELTIETDILEKKEI 1362
QY 1307 KPEQVSEGEQTAAPHEGTYGKPVLTLDMPSE----- 1341
Db 1363 EKDHFEKEFEAEIADLEADILKEVSSLEVEBEKLEEVHELKEVEHIIISGDAHIKGL 1422
QY 1342 -----RGKALGSLGSPSLPQDKAGCIEVQVQSLDVTVTQTAABAVEKVIETVV 1390
Db 1423 EEDDLEEVDDLKGSILMDLKGDMELGDMDK-----ESLEDVTTKLGRVES-LKXVL 1473
QY 1391 ISETGESPECVGAH-----LL-----PAEKSSATGHWTLQHAEE--DTVPLG 1430
Db 1474 SSALGMDSEQMKTRKKAQRPKLEVLKEVEKPEKKIITKKVRFDIKDKEPKDEI-VE 1532
QY 1431 PESQAESIPITVPAPESTLHPDLQGEISASQSRSEERSEEDPKDPAGPDAGKSTAEKVL 1490
Db 1533 VEMKDEDI-----EEDVEEDIEEDKVEDIDEDIDEDKDKD-----EVIDLIV 1580
QY 1491 KAPEILELESKNKI 1506
Db 1581 QKEKRIEKVAKKKKL 1596

RESULT 5
US-09-864-761-34248
; Sequence 34248, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34248
; LENGTH: 2665
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL034555.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 12
; OTHER INFORMATION: EST_HUMAN HIT: AU117052.1, EVALUE 0.00e-00
; OTHER INFORMATION: SWISSPROT HIT: P08640, EVALUE 3.00e-10
US-09-864-761-34248

Query Match 5.3%; Score 429; DB 10; Length 2665;
Best Local Similarity 19.6%; Pred. No. 2,4e-11;
Matches 372; Conservative 270; Mismatches 755; Indels 504; Gaps 83;
QY 9 QRSPEQAGSD-----TPSELVLSCHGHPAAASG-----AAGDPADADPAT- 49
Db 301 QVSTEPAKSDLSKLSVRMKVPKEKGLSHVVEVEKEGLKARKHLKPEQPDGVSVD 360
QY 50 --KLQKNGQLSSVNGVAGQGVHVOENQOEVEEVEDVQGRESEDVREKDRVEEMA 107
Db 361 LEKLEARKRRFADSNLKAQKQPEVKKSPMEDARVL-----SKQPDVSSREVI- 411
QY 108 ANSTAVEDITKQOEETSIIEIQIPASENNVEEMVQPAESQANDVGFKKVFGFKFTV 167
Db 412 -----LLREGEARKPKVRKEILKRES-----KKI----- 435


```

QY 168 KKDNEKSDTVQVLLTVKKDBEGBAEASVAGDHPESVETAVGESAASELQSTKOE 227
Db 436 -----KIDRLINTVASPKDCQELASISVSGSGRPSSDLOARLIGELAG-----ESVENOE 483
QY 228 GTLKQEOSTELP-----LOAESDOAAEEB--AKDGBEEKQEKEPKSPSPSPNS 278
Db 484 -----VQSKKPLPSKPPOLKQLOVLDQGPREDVRKNKICSLDELTPERKSGQEKSHSVNT 538
QY 279 ET-----TSSFKKFTTHGAWGRKKTSPFKSK-----EDDLTAEKR----- 315
Db 539 BEKIDIDIDHQSRYKQWQESRRKQWMEIASEKFGSPKQDVVEYERRSLVHEGKPP 598
QY 316 -----KEQAEKVD-----EBEKEKTEPASEQEPAEEDTDQARLSADVEKEL 358
Db 599 QDVTDSPSPSKKRMDDHVDICTRKERNYRSRQISBESERTGSPSVRHSGFHEH--- 655
QY 359 PLEDOVGLEASS-----BEKAP-----LATEVFDEKMEAHQEVAVHVS 401
Db 656 --EDYIGSPRLISVKGSPKVDKVLPSYNTTYRESLKNPFDSSRRQOMADMAKIKLSV 713
QY 402 VEKTEEOGGGGEAE--GGVVEGTGESLPPEKL-----AEQEVPOEAPAE-- 448
Db 714 LNSEBELNRWDQMQODARFVSPNSIIRKDSLKRKSVRLDEPGEVPSDEDEGHEKS 773
QY 449 -----LMKSRMCVSGGDHTQLTLSPE-----EKLPHKPEGIYSE 485
Db 774 HSPRASALYESRLSFLNRREDKLERDERLSSLSERNKFPYFALDKITPTDTALER 833
QY 486 VEMLSS-----QERIVQGP-----LKKLFSSGKKXLSG 516
Db 834 AKSLSSSREENWSFLDWDSRFANFRNNKDKEXYDSAPRPIPSWYMKKKIKRIDSQKMD 893
QY 517 KQOKKRGGGGDEEGEYOHITHESPESADEQKGSASSPEPEPTTCKLEKPLAPOD 576
Db 894 KKEDHK-----EEQEROLFLASRFLHSSIFPDQSKRLQHLERKEDSPFISGRYKQT 948
QY 577 GEABEGTSDGKKEGTPMAS-FKKV-----TPKKVRPSPESDKELEKVKASIS 631
Db 949 SEGANSTT--DSIOEPVLFHSRFMELTRMOQKEKQKKEVEKQEDTNHPTPES 1005
QY 632 STDSTVSEWQ-----DEVKTVEGQKPEPKRVDTSVSEWALICVSSSKKARKAS 684
Db 1006 AEPNDSELTPTPSVGPSTVVTLESAPSALEKTTGKTVEAPLV--TEETVYEPATV 1062
QY 685 SPD-----EGPRTLGGDSHRAE-----ASKDEAGTDAVASTODQOAGSSPEP 733
Db 1063 SEBAKRPASPAAPV-----EQLEQVLDLPPGADPDKEAA--MMPAGVE-----GSGGDQP 1111
QY 734 ---AGSPSEGGVSTWE-----SPKRLVTPRKKSQKLEKAE--DSSVE----- 773
Db 1112 PYLDAPPTPGASFQAEENVDPEPDSTPLSKPAQKSEANEPAKEDATADAPDAN 1171
QY 774 ---QSTELPEPREEBWSIKKFIPIGRKKRADGQEOQTVESGVEINEDPNVPAV 830
Db 1172 OKAEAPESQPPASED-LEVDPPVAAKDKPKNSKSKTPVQ-AAAVSVI--KPVT 1224
QY 831 PLEEVAVAREKMEAOGNT--ELPOLLAGVYVSELSKT-----LVHTVSAVIDGT 880
Db 1225 RKSF--RIRREKUKRNSRGEQKLEKMEKIKTRTASKNASADLHPESPLSKT 1282
QY 881 RA--VTSV-----EERSPSWISASVTEPLEHTAGEAMPVEEVEKQILAE--ETPV 929
Db 1283 RRRNVSVATWGDHENSPP-----VKEPVEQ-----PRVTRKLERELQAAAV 1327
QY 930 TOT-----LPEGDADHDWTSVVDTSSEAVNTIETSEALRTVEVEAGAE--ETTD 980
Db 1328 PTPPRGRPPKRRRRADBEENAKAPAEFTLKPBEGRMSRSPQKTAAGGPOQKKNP 1387
QY 981 MYSASQULDSPTTEAPAPVQGVESGVLDTEEBEROTQIILQAVADKYVEEQVATQT 1040
Db 1388 KVDAT-----REPAITVEGPQIGVKESSMPEKAAEBEAGEQKRDKDACTDNKPTIAP 1442

```

```

QY 1041 VORTGKALEKVEEVEDS-----EVLASEKED--VMPKGPVOBAGAEHL 1084
Db 1443 VE-----VEKKAPAEKNSKSGKGRNSRLAYDKASLSLKNVDAASPRGAAGAGER-- 1495
QY 1085 AOGSETG--QATPESLVEPVTADVDHATQVYIKQOLMEQAVAPESSETLTDSETNGS 1142
Db 1496 -----ESGVAVVSPKESSPQKEDOLS-----SQLKSDPVDPKPEKEKEDVASAGP 1541
QY 1143 TPLADSTAGTQODETIDS-----QSKATAAVRQS-----QYTEEAAT 1183
Db 1542 SPEA--TQIARQWELQAVEHIAKLAESASAAAYKADAPBGLAPEDRDKPAHQASSETLAA 1600
QY 1184 AQK---EESTLPPNNVPAQEBH--GE-----PGRDVLPTQOEL--TAAAVPLAKTEV 1231
Db 1601 AIGSTINDISEPNEFPAPPPYFESQTDLOPPAGNALQPSSEGMETDDEAVSSILETFA 1660
QY 1232 QOGEVDMLDGEKYEQEVFVHSGPNSQKADVTYDSEVWVAGCQEKESTEVOSLSLE 1291
Db 1661 ATESSRPVNAPODS-----AGPTDTKEARGNSSETSHSVEAKSGKEVEVTLVRKD 1712
QY 1292 EGMETVYEKREKTEKQYSEBEOETAPENHGTGKXVLTIDMSSSRGKALSG 1351
Db 1713 KGRQKTRSRKENTNKVVA--PVESHVESNOAG--ESPANEGTTV--Q 1759
QY 1352 SPSLPDQKAGCIEVOYOSLDTVTVOTA-----BAVEKVIETVVISSETGESPECVGAH 1404
Db 1760 HPEAPQEBK-----QSEKPHSTPQOSTSLKIPSTENSSQGISVEBKRTPTASVPPD 1813
QY 1405 L-----LPAEKSSATGHWTLQHAEDTVPLGPGSAESIPPIITPA----- 1445
Db 1814 LPPPOPARVDEBQAPRVHISIESDPVTPPD-PSIPITPLSVTAALKSPVVASGI 1872
QY 1446 PESTHPLQGEISASQRENSSEEDKPDAGDADGKE---STAIKYLAEPILELES 1501
Db 1873 PHQSPPTKVTMTIROBEPRAQSTPSPALPBDTKASVDTSSTLRKITL-MDPYVATS 1931
QY 1502 KSNKVLNVITQAVDQFARTETAPETHAYDSQTOVPACRDSREPNCWTGKAKMKHP 1561
Db 1932 VTSISVTTALAEV-----SABCLH-----EAPPEPVDSKAP-----LEEKTAAP 1972
QY 1562 VPOREDLQVLTLEAMQPRKCLPRLQK-----APVS 1595
Db 1973 VTN-NSEIQASEVLVA-ADKEKVAPVIAKITSVISRMPVS 2011

RESULT 6
US-09-815-242-5635
; Sequence 5635, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5635

; LENGTH: 2368

; TYPE: PR1

; ORGANISM: Staphylococcus aureus

US-09-815-242-5635

Query Match 5.0%; Score 407.5; DB 10; Length 2368;

Best Local Similarity 17.5%; Pred. No. 1.8e-10;

Matches 317; Conservative 293; Mismatches 663; Indels 541; Gaps 71;

QY 47 PATKLPKNGQLSVNGVAEGDVHVGEOHQEE--EVVDVEDVQRESDVREKD--- 101

DB 677 PTVKPQAKQDIQAVTRKQ---IKKSNASLQDEKDVANDKIGKJETKAIDRAATT 732

QY 102 RVEMAANSTAVEDITKDGQ-----BETSEI---IEQIPASNNVEMVQPAESQ 148

DB 733 NAQVEAIKTKAINDINQATPATTAACAAALEEFDEWVQAQIDQAPLNPDTTNEEVAEAIER 792

QY 149 AN--DVGFKKVFKVPGFKFTVKDKKNEKSDTVQLLT---VKKDEGEGEASVGCAGHGE 202

DB 793 INAAKSGVKAIEATTTAQDLERVKNEEISKIENITDSTQTKMDAYNEVKQAATARKTON 852

QY 203 PSVETAVGESAKSELKQSTEK-----OEGTLKQBSSTE-----IPLQAESDQ 247

DB 853 ATVSNATNEEVAEADAEEAAQKQGLHDIQVVKSKQEVADTKSKVLKINAIQTOAKVKP 912

QY 248 AABEEAKDE--GEEKQEKPTKSPSPSPVNSSETTSSFKFFFTGHWGRKKTTSFKKSK- 305

DB 913 AADTEVENAYNTRKQEIQNS-----NASTTEEKQAAAT-----ELDTKKQEA 954

QY 306 -----EDDLTAR-----KKKEQEAKEVDEEKEKT-----PPAS 335

DB 955 RTNLDAANTSDVTTAKDNGIAAINOVQAAATTKSDAKAEIAQKASERKTAIEAMNDSTT 1014

QY 336 BEQBPAD--TDQARLSADYKVELPLEDQVGDLEASSEKCA-----PLATEVFDE 385

DB 1015 EQQAQKQVDQAVVTANADIDNAANTVDVNAKTITNEATIAITPDANVKPTAKQAIAD 1074

QY 386 KMEAHQBVAAEVHSTVTEKTEEEQGGGEABGGVVVVGTCGESLPPKLABPQEVPOBAEP 445

DB 1075 KVOAQETAI-----DANNGATTE-----EKAAAKQVQTEKTT 1107

QY 446 ABELKSRMCVSGDHTQLTDLSPBEKTLPKHPEGIVSEVEMLSOERIKVQSGPLKLL 505

DB 1108 ADTAI-----DGAHTN-----AEVEAAKNAEIAKIEAIQPAATT 1140

QY 506 FSSSGLKLGKQKQKGGGDEEPGEYQIHTE-----S 541

DB 1141 TKDNAKQAIATKANERKTA-----IAQTDITAEETAAANANVDNAVTOANNIEAANS 1194

QY 542 PESADEQK--GESSAS--SPBEPBETCL--EKGPLAPQDGEAEGTSDGCKKREGIT 595

DB 1195 QNDVDQAKTTGEASIDQVTPVVKKATAVTDKNNITAAATDNDGVDTKADGAKNSIGSTQ 1254

QY 596 PWASFKK-----MVTPKKRVRRPSESDEE----- 620

DB 1255 PATAVKSNAKNDVDQAVTTQNAQIDNTTGATTEBKNAAKDLVLKAKEKAVQDILNAQTN 1314

QY 621 ELEKVKSATLS-----STDSTVSEM--ODEVKTVCGEOK-----PEEPKPRVD 661

DB 1315 DVTOIKQOAVDVGIGTADTTIKDVAKDELATKAREOKALIAQFADATTEKEQANQOVD 1374

QY 662 TSVSW-----BALICVSGSKKRKA-----SSDDEGGPRT-----LG 695

DB 1375 AELTQGNQNIENAGSIDDVNTAKNAQAIQIDPIQASTDVKTNARAEILLTEMONKITEILN 1434

QY 696 GDShRAEASKO-----KEAGTDVAPAST-----QEQDQAQGGSSSPPEPAG 735

RESULT 7

US-09-815-242-12389

; Sequence 12389, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

```

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 12389
LENGTH: 2368
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12389

```

Query Match 5.0%; Score 407.5; DB 10; Length 2368;

Best Local Similarity 17.5%; Pred. No. 1.8e-10; Indels 541; Gaps 71;

Matches 317; Conservative 293; Mismatches 663;

```

47 PATKIPKQGLSVNGVAGDGVHVOENOGQEE-BVVDVGVGQRESDVREKD--- 101
Db 677 PIVKQAKQDIIQAVTTRKQ---IKSNASIQDEKDVANIKIGIKETKAKIDDAATT 732
Qy 102 RVEEMANSTAVEDITKQGO---ETSET---TEQIPASENNVEEMVQPAEQ 148
Db 723 NAQVBAIKTKAINDINQTPATTAKAALAEFEDEVQAOIDQAPLMPDTTNEVEVAEIR 792
Qy 149 AN--DVGFKVKKFVGKFTVKKKDKKREKSDTVQQLT---VKKDEGEAGAVGAGDHQE 202
Db 793 INAAKSGVKALEATTTAODLERVKNEBLSKIENTIDSTQTKADAVNEVKAATPAKTION 852
Qy 203 PSVETAVGESAKESSELKOSTEK---QEGTLKQEOSSTE---IPLQAESDQ 247
Db 853 AIVSNATNEVEVAEADAABAQOGHLDQVVKSKQEVADTSKVLKINALIQTQAKVXP 912
Qy 248 AAEBAKDE-GEKQKEPTKSPESPSPVNSETTSPFKKFTTHGAKRKKTSPFKSK- 305
Db 913 AADTEVENAYNRKQEIOMS---NASTEKEQAAYT---ELDTKQKOA 954
Qy 306 -----EDLETFE-----KRKEQAEKVDKEEKEKT-----EPAS 335
Db 955 RTNLDAANTNSVTTAKGNGAIAINOVQAATTKSDAKAEIAOKASERKTAIEAMNDSTT 1014
Qy 336 EEOCPAED-TDOARLSADYEKVELPLEDQVGLDEASSEKCA-----PLATEVPDE 385
Db 1015 EEOQAAKQKVDQAVTANMDINDAANTVDNAKTNEFTIAITPDANVKPTAKQALMD 1074
Qy 386 KMEAQOEVAEAVHSTVEKTEEBEGGGGEAEAGGVVEGTGESI.PPEKLA.PQEVPOEAP 445
Db 1075 KVOAQETAI-----DANNGATTE-----EKAAKQOVQTEKTT 1107
Qy 446 AEELKSRMCVSGDHTQUTLDSPEEKLPKHPBGIVSEVENTLSSQDEIKYQGSLLKTL 505

```

```

Db 1108 ADTAI-----DGAHTN-----AEVEAAKNAIEAKIAIQPATI 1140
Qy 506 FSSGGLKLSGKKQKGRGGGDEPGEYQHTE-----S 541
Db 1141 TKDAKQAIATKAEERKTA-----IAQTQDITAEERIAAANAANVANAQTQANNIEANS 1194
Qy 542 PESADEOK--GESSAS--SPEEPTTCL--EKGPLEAPQDGEAEGETTSDEKKRGIT 595
Db 1195 QNDVDQAKTTGEASIDQVTPVVKKATAVVDAKNNITTAIDNDGVDTAKDAGKASISQIQ 1254
Qy 596 PMASFKK-----WTPPKRVRRPESPDKE----- 620
Db 1255 PATAVKSNAKNDVQAVTTQQAIDNTTGATTEKNAKDLVLAKKXKAYQDILNAQTIN 1314
Qy 621 ELEKKSATLS-----STDSTVSEM--ODEYKTVGEBOK-----PEEPKRRVD 661
Db 1315 DVTOIKQAVADVQGITADDTIKDVADELTAKREKALIAQTADATTEKEQANOQVD 1374
Qy 662 TSVSW-----EALICVGSKKRKARKA-----SSDDEGGPRT-----LG 695
Db 1375 AELTQGNONITNAQSIDDVNTAKDNAIQALDPIQASTDVKTNAABELLTEKQKITELIN 1434
Qy 696 GDSHRAEERASXD-----KEAGTDAVPAST-----QEDDQAGSSSPERPAG 735
Db 1435 NNETTNEKNDIGPVRAAYEEGLNINIAATTGGDTVTAQTAVQKQQLHANPVKKPAG 1494
Qy 736 SPSEGEVSTWSEFRKLVTPPKKSKSKLEEKAE--DSSVEQLSTIEPSPRESWV----- 788
Db 1495 KTAIDQAAAD-KKQIQIEQTPMASQOEINDAKQOEVDTEINQAKTINIDOSTDEYVDNAVKE 1553
Qy 789 -----SIKKFIPOGR-----KKRADG 804
Db 1554 GKAKINAVKTFSEYKQALAKIEAIVNAKTEADNSNASTSEETAEAKQKALBKQTDQ 1613
Qy 805 KOEQATVEDSGPVEINEDDPNV-----PAVPLSEYNAVEREKMEAOGNTLEPQ 853
Db 1614 NVNQATSKDIDLEVOIHNDLDINDYITPTGKKEASATDLVAVADQKNNISADINATQDE 1673
Qy 854 LLGAVYSEELSKTILVHTVS-----VAVIDGTRAVTSYEERPSMISASV 898
Db 1674 KQQAIAKQDVQVQTALESINNGVNDGVDALITQKKAIDAIQVADATVPKPAANOVIDAKA 1733
Qy 899 TEPLP-----HTAGEAMPVEEYTE--KDIIAEFTV--LTQTLPGKGAHDNV 944
Db 1734 EETKESIDQSQLTAETKTEKLAMIKQITDQAKGQITDAITTAVEREAKKAGLAPDIT- 1792
Qy 945 TSEVDFTEAVTATETSEALTEETVEASGAETTD-----MVASQQLTSP 992
Db 1793 --QIDSTKQKAIBELETAALQIEAGVNVADATTEBEKAFTNLLEDLSKATDISQIT 1850
Qy 993 DTTEBATPVQEVESGLDTEEBERQTAI---LOAVADKYKEESQV-----PATQTV 1041
Db 1851 TNAEIAAT---VKNSALEQLAKRINPVKKNALEAIREVNNKQIEILIKNADADASAKEI 1906
Qy 1042 QRTG-----SKALEKVEVEEDSEVLASEKEKDVMPK---GPQGEAGAEHLAAGSE--- 1089
Db 1907 ARTDLGRYFDPADLDLTQNTTEV--AELQNVTTIPALIEAIVPONDPNANDTNSGSDND 1964
Qy 1090 -----TGOATPESLEVPETADVVDVATQVYIKLQULMEQAVABESSSETLTDSETNGSTP 1144
Db 1965 ATANSNANATPENTQCPVHTESTDANAN-----DTSGTTNNQNDATG 2008
Qy 1145 LADSDTADGTQOQDETID--SQDSKATAAVROSQVTEBEAATAQKEEPTLPNN---VPA 1198
Db 2009 ETTATSANSATDANDKXPQANNNSADTSTNSPTMDVTSKPEVEST--NGTTDPA 2066
Qy 1139 QEHEGEEGRVLE-----PTQOELTAAAVVLAKTBVGQGEVDMIDGEXVK 1246
Db 2067 TEADNATPAESATNNSTTTTATNENAPTESPATY--PTTAST--GAESSADSKONASVN 2121
Qy 1247 E-EQEVFVHSG-----PNSQKADVTYDSEVGVAGCOEKESTVQSLSLEG 1293
Db 2122 DSKQNAEYNNASQSTNGKTAQPKSENKAKAEKDRGDRSTQSWESTTETLPASADITEP 2181

```

Qy	1294	EMETDVEKEKRETPQGVSE---	GGQETAAPPEHGTG----	KPVLTLDPSSER	1342
		: : : :	: : : :	: : : :	
Db	2182	KVSNSTKDKKEESTTSTQDAEQHNSDTNVASNEAKSGE	NDTDSVKP--	STGKPSSEAK	2239
		: : : :	: : : :	: : : :	
Qy	1343	GKALGSLGGSPSLPDODKACGIEVQVSLDITVTQTAAE	VKVIETTVISGTSPECVG	1402	
		: : : :	: : : :	: : : :	
Db	2240	DKA-----	TSTEDSQKDMATADTKNQASIGATADVNNKATON-----	DG	2280
		: : : :	: : : :	: : : :	
Qy	1403	AHLHPAEKSSATGGHWTLQHAETDPLVGPES-QAES	PIPIVTPAPESTLHPDLQGEISAS	1461	
		: : : :	: : : :	: : : :	
Db	2281	ANASPA-----	TVSKGNSANQDMLNVTKTKENAKAKSAQQQKVN--	2321	
		: : : :	: : : :	: : : :	
Qy	1462	QRESSEEDKPDAG	1475		
		: : : :	: : : :	: : : :	
Db	2322	-KPKQAAKTLPDG	2334		
		: : : :	: : : :	: : : :	

```

RESULT 8
US-09-864-761-35612
; Sequence 35612, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35612
; LENGTH: 1325
; TYPE: PRT

```

; ORGANISM: Homo sapiens

```

FEATURE:
; OTHER INFORMATION: MAP TO AP000511.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: AW867076.1, EVALUATE 1.00e-103
; OTHER INFORMATION: EST HUMAN HIT: AW867076.1, EVALUATE 1.00e-103
; OTHER INFORMATION: SWISSPROT HIT: Q02817, EVALUATE 5.00e-63
US-09-864-761-35612

```

Query Match	4.8%	Score 388	DB 10	Length 1325
Best Local Similarity	20.3%	Pred. No. 6.6e-10		
Matches 326	Conservative 191	Mismatches 662	Indels 428	Gaps 68
QY	53	QKNGQLSSVNGV-----	-----AEQGDVHVQEEN-QEGSEEV	84
DB	8	QKTGEIUSTSDHIPLPLGLVYSIPFDHIVLHSGORPPELPSKTEIHQKHCNTRISKP	67	
QY	85	VDEDVQORESEDVREKDRVBEEMAANSTAVEDITKXGOEETSEIIEQIPASENNYEEVMQV	144	
DB	68	TDKPTG-NSKTIIDHKSSTDNHEAPPTSEENSSNOGKD-----PMIRN--QRSVDP	114	
QY	145	AES---QANDVGPKVPKFGVGFETVKKDKNEKSDTVQLLTVKKDEGEAGASVAGDQH	201	
DB	115	ADSTTTKESAGKHHITPAPKINCKRSTTGKS-----TVTRKSDK-----TGRLP	161	
QY	202	EPSVETAVGSASKESELKOSTEQEQTLEKQOSSTEIPLOAESDQAAEEAKDEGEKQ	261	
DB	162	EKMSIL--DKTSTSSHKTITSFHNSGNSQTQKSTFP-----EKITAASKT	207	
QY	262	EKPPTKSPESPSPVSNSETTSPKFPPTHGAWGRKKTSPKKSKEDDLETAERKKEQAE	321	
DB	208	TYKTTGTPEBSEKTEDSRITVASDKLLT-----KTKTKTQETISANELTQSL	254	
QY	322	KVDEEKEKTEPASEOEPAEDTDQARLSADYEKVELPLEDQVGDLSASEEKCAPLATE	381	
DB	255	AETEHGGRTPANENNTTSPPAEPTENRERTAN-ENTTLPSPAEPTENRERTANENTAP	309	
QY	382	VFDEKMEAHQEVVAHVHVSTVEKTEEBEQGGGEABGVGVVGGTGESLPPEKLAEPQEVQP	441	
DB	310	-FPAGPTENREMTANENTTLFPAEPTEHG-----ERTANENTTSPS	348	
QY	442	EAPPAEELMKREMCVSGGDHITQLTDSLPEKTLPKHPEGIVSEVEMLSQERIKVQSGP	501	
DB	349	PAEPTEH-----GERT-----ANENTTP-----SPAEPTEHGERTPFAND-	383	
QY	502	LKLLFSSSGLKLSGKKQKRGGGDEBEPGEYQHITHESPEGADEQKGSASSPPEPE	561	
DB	384	--KTSSS-----AESTEHGERTPPLANENTTSPPAEPTENRERTANENTTSPPAEPTENR	436	
QY	562	ETTCLEK--GPLEAPQDGBA---BEGTTSDGKKREG-ITPNASFKKOMVTPKKVRRRPS	614	
DB	437	ETTANKEKTLTSPVEPTENRRETTANKEKTTSPSPAPTENGORTTFFANEK-----TTSS	487	
QY	615	ESDKBELEKVKATLSSTSDTSVSEWQDEVKTVGEB--QKPEPKRRVDTSVSWEALICV	672	
DB	488	SABPTTEHGERTPPLANENTTSPPAEPTENRERTANKEKTTSPPAEPTENGDRT-----	538	
QY	673	GSSKKARKASSDDGGPRT-----LGGDSHRAEASDKKEAGTDAVPASTOEQD	723	
DB	539	PLANENTTSPLAPTENGORTTFFANEKTTSSSAEPTHEERTPLANENTTSPPAEPTENR	598	
QY	724	Q--AQCSSPPEPAGSPSEGCVSTWESFKLVTPRKKSCKLEAKEDSSVEQLSTIEP	781	
DB	599	ERTANENTTSPAG-PTENREMTANE--KTTLPFAEPTENR-ERTANEKTT--SSPAEP	651	
QY	782	SREESWVSIKFFIPGRKKRDKQEOATVEDSGPVEINDDPNVAVPLSEYNAVE-R	840	

```
Db 652 TEN-----GQRPFPAN--EKTSSPAEPTGEKRTPLANENTTSSPAEPTENR 697
Qy 841 EKMEAGCNTELPOLLGAVYVEBELSKTLVHTVSVAVIDGTAVTSVEBRSPWSIASYTE 900
Db 698 ERTANEKTTTPP-----AEPTENRERTANENTTP-----SPAQ 730
Qy 901 PLEHTAGAMPPEVEVEKDI1AETPVLTQTLPEKGAHDMTSEVDFSEAVTALET 960
Db 731 PTEN--GDRTPLANEKT-----TPSLAEPTENGK-----RTPFANEKTT---S 768
Qy 961 SEARTEEVTEASGAETTDVMSAVSQTLDSPDTEEAATPVQEVESGVLDTSEEROTQA 1020
Db 769 SSAEPTHEAERTPLANENT-----TSSP-----AEP-----TENRRTAN- 803
Qy 1021 ILQAVADVKEESQVAPATQVORTSKALEKV---EEVEEDSEVLASEKEKDVMPKCPV 1076
Db 804 -----EKTQFPAPTEPRE--STANEKUTPPAPTEPREMTANE--NTTISPAEPT 852
Qy 1077 QGAGAEHL-----AQSGETGQATPESLEVEPVTVADVHVATCOYIKLOQLMEQAVA 1127
Db 853 EHEEMTPLANEKTTLSPAPTEPTENGERTP-----FTNEKTT 887
Qy 1128 PESSRTLDSENGSTPLADSP-----ADGTQDETTIDSODSKATAA--VROSOVTEEA 1181
Db 888 PSSAA--PTEHEKRTPLANEITTPSRABPTHEGERIANEKAETSPAPKTEHEGTTVND 944
Qy 1182 ATAQKEEPTLPNNVP--AOE-----EHGEERGRDVELEPTOELTLAAVPLAKTE 1230
Db 945 TTPSAEFTENGERTPLANENTTSTPSTEHGERTANEKTTSPAPTE----- 994
Qy 1231 VGOGEVDMLDEKYEEOEYFVHSGPNSOKADVTYDSEVWGAVGCOEKESTEVOQLSL 1290
Db 995 -----HGERTPSANEKTIPSPAPKTEHEMTPPSA-----NENTTPSPVKP 1034
Qy 1231 EREGMETVEKEKRTKQOVSEEGEOETAPEH--EGTKGFVLTLDMSSRGKALGSL 1349
Db 1035 TEHGKTLTLANEKTILSPBEGPTEHAKTTISANEKTTPLAKTEHGERTSINDKITS- 1093
Qy 1350 GGSPLPQDQKACIEVOVQSLDTTV---TOTAAVEKVIETVVISSETGESPEVCAGHL 1406
Db 1094 -AAESTERDRATSANVITTPAPABPIKAKRTTLAHEKK--TOVTEKSTHEKETS-- 1147
Qy 1407 PAEKSSATGHWTLQHAEDTV-----PLGRESQAESt-----PIITTPAPE 1447
Db 1148 TTEKTRTPREKPTL-YSEKTLCTKGNKTPV--PEKPTENIGNTLTLETIKAKVKSSTENPE 1205
Qy 1448 STLHDDLQGEIISAGRESEEDKPDADGKSTALEKVLKAPETILEESKNKIY 1507
Db 1206 KT-----AAVTKIKPSVKYTGKSLTITSSHLNKE--VTHQVPTGSFT 1248
Qy 1508 INVIGTAVDOFARTEAPETHAY---DSOTQVPACRLDLSREPNRCW 1550
Db 1249 LITSKTKLSITSEATGNESHPLYLNKDSQKGIHAGQEMENDSPFAM 1295

RESULT 9
US-09-971-536-70
; Sequence 70, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilka
; APPLICANT: Bloksberg, Leonard
; APPLICANT: Lubbers, Mark
; APPLICANT: Dekker, James
; APPLICANT: Christenson, Anna
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul
; APPLICANT: Reid, Julian
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Lactobacillus rhamnosus polynucleotides, polypeptides and Methods
; TITLE OF INVENTION: Using them
```

```
FILE REFERENCE: 104362
; CURRENT APPLICATION NUMBER: US/09/971, 536
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: PCT/NZ01/00160
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 1879
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-971-536-70

Query Match 4.7%; Score 382; DB 9; Length 1879;
Best Local Similarity 19.1%; Pred. No. 1,8e-09;
Matches 324; Conservative 302; Mismatches 726; Indels 342; Gaps 70;

Qy 20 TPSELVLSGHGPAASGAAGDPADAPATKLPG-KNGQLSVNGVABQG--DVHVQOE 75
Db 124 TATEKAKQKQDVAAEATKAKAIDQAGTEGVOQADAGIAIDQHQPGTGLNREBEA 183
Qy 76 NQGOEEV---VDEDVGRSEEDVREKDRVEEMANS-TAVEDITTKQGEETSEIIE 129
Db 184 KQAIIDAEAAKVTAEIEDSTLATEKKAQKQGVADAEAKAKATAID-----QAGTIEAID 237
Qy 130 Q-----IPASENVVEEV--OPAESQANDVGFKKVFX 159
Db 238 KAKDQGIKAIIDQKQADPPTRAQAKAIDAEAAKKAIDDPILTAKDXTAKQ--Q 295
Qy 160 FVGRFFTVKQDKNEKSDPTQLLVTKDEG--EGAEASVAGAGH---QEPSEYAVGESASK 215
Db 296 GVGBEATKAKTAIDQAKTIDGVIQAKDGIKAIIDQHQAGTDLATRRDSAKQAIIDAEAK 355
Qy 216 -----ESELKQSTEKO-----EGTLKQEOSSEIPLQABSDQAAEEBAKDE 256
Db 356 ITDAINDDTLTSTKQKQKQAVADAEAAKKAIDQKQNDADAI--LQAOADGIKAIIDAHQ 414
Qy 257 GEEKQKEPTSPSPSPSVNSETTSFKPFTGMAWGRKKTSTFKSKEDPLETAERK 316
Db 415 IGADLDYKTKAKQA-----IDKEAA--KVL-----AIEDPILTSAEK 453
Qy 317 EQAEKVDDEEKEKTEPASEQEPEDTDQ--ARLSAD-----YEKVELPLEDQVGL 368
Db 454 AQKQGVADETAKAKTAIDS-----ARNADELAKAQADQIKAIIDQHRLGMDLARKTDQ 508
Qy 369 ASSEKCAPLATEVFDKMEAHQEVVAEIVSTYEKTEEEO--GGGGEABEGVVE--- 422
Db 509 AAIDAEAAKVEAIDDPPTLTSQEKAAQKQTPAAEATKAKATIAKADADVIDIAEKAGI 568
Qy 423 -----GTGSLPPEKLAEPQEVPOEAPAEELKMSKEMCVSGDHQ---LTDLSPEE 472
Db 569 QAIIDGHQSGALLDTRKVDAKKAIIDAEAKTNDADIDDPVLTSAEKATQKQKVTDEAKA 628
Qy 473 KTLPKHREGIASEVEMLSSQERIVQSGPLKFLPSSGKLKLSGKQKGRGGGDEEPG 532
Db 629 KT-----AIDAKNMDVTDQAKASGIGALIDAVHQS--TLIDTRKQDAK--AIDAEAV 678
Qy 533 EY-----OHHTESPESADEQKSGSSASPEPEPTTLEKGPLADPODGAEBEGTSSDG 587
Db 679 KVIAAIGQDVTLTQAEKTLTQQAADAVDATQ-----AKAAIDAKNMDAVDQAKDG 729
Qy 588 EK-----KREGITPMASRKKKVTPKRRVRPESDKEELEKVSAA-----TLSTSTV 637
Db 730 IKAIDAQHQAGL--ALMERKAANKL-----AETADKVDQAIGQDVTLTATQAV 778
Qy 638 SEMODEVKTVEEQ-----KPEPRRRVDTSVSWBALICVSSKKRRARAS 683
Db 779 QRQAITVEVTRANQALDAAGNADAVDQAKNKGVAIYDQHQSGAL-----ADRRKDAK-Q 833
```

```
Qy 684 SSDDEGGPRTLGDSH---RAEASKQKEAGTDAVPASTQEQDQAQSSSPPEPAGSPSE 740
Db 834 AIDAEAKETAADODATLTANAKASQQAQVADATKAKEAIDAAKQADAVDQA-----KN 889
Qy 741 EGVSTWES---FKRLVTPRKSKSLEKAKSDSSVEQLSTEIEPSRESWSWIKKFTIPGR 797
Db 890 DGIRAIQAQHHAGQAVADRKA---AKAIDAERAKVTGNID--QDETLDATK-----938
Qy 798 RKKRADQKQEQATVEDSGFVINEEDPNPVA--VPLSEYNVAREKWEAQG-----NTEL 851
Db 939 ----AAQQAQVATEADNAKQAIDKQ--NADAVDKAKTGKGIKAIDAQHQSGQAIRQONDA 993
Qy 852 POLLAGVTVSEELSKTLVHTVSVAVID---GTRAVTSVEERSPWSIAS--VTEPLEHTAG 907
Db 994 KQALDA--EAAKVTAIDQDDPTLTAEEKKAQQAQVATDAETKAKAIDATLVADAIQALA 1051
Qy 908 EAMPVPEVEITEKDI--IAEETPVLTQTLPEGRKADHDMVTSEVDTFSEAVTATETSEALRT 966
Db 1052 DGIKTIDAQYQTGIALDKKAAKQTI-----DAEAAKVSEAIID--QDVTLTADQKATQK 1104
Qy 967 EEVTEASGAETTMVSAVSLTDSPTTE-----EATPVQEVESGVLDTTEBERQTA 1020
Db 1105 QAV-----ADEATKAKAAIDQASDADAVIQATIDGIEAIDAQHQSAVALD--KQKQQAQK 1157
Qy 1021 ILQAVADKVEESQVPATQTVORTGSKALEKVEVEEDSEVL-----ASEKEKDVMP 1072
Db 1158 AIDAEAAKSKAIDODVTLTATQKADQQAQVIAEADKAKKLIDAGNADGTKQAESDGIK 1217
Qy 1073 KGPVQEAHGAELAQSGSETGQATPESLEVPETADVHVATQVVIKQQLMEQAQVAPESSE 1132
Db 1218 AIDAQHQSSQALADRKDAK--TAIDAEAAKETAIDHDATLTA-----NEKA---SOKQ 1267
Qy 1133 TLTDSETNGSPPLADSPDADTQOQDET-----IDSQ-----DSKATAAVRQSQVTEE 1179
Db 1268 AVTDEATKAKKAIDAAKQAQADAVDQAKTDGKAIKIDAQHHSGQALDRKADA---KQVIDA 1323
Qy 1180 EAA--TAQKEPSTLPNNVPAQEEHGEPEGDRDVLPTQOQELTAAAVPVLAKTEVQGEV 1237
Db 1324 EAAKVTAIDQDNTLTAKQAQAQKQGVATEAD-----KAKQAIDAAGDA 1367
Qy 1238 DMLDGEKVEKEQEVFVHSGPNKQAADVTYDSEVMGVAGQEKESTEVQSLSLEEGMET 1297
Db 1368 DAVDQAKTAGIAI-----DAQHAGKTIIDSRHDDAKQAIDEEAAKV-----IKAIODDP 1417
Qy 1298 DVEKERETKPEQVSEGEQETAAPHEG-----TYG-----KPVLTLDMPSS 1340
Db 1418 TLTAQAQEAQQAQVATEADKAKKAIDAAQADAVDQAKTAGIKALDEQHKSGQIVDARKE 1477
Qy 1341 ERGKALGSLGSSPLPDQDKAGCIEVQVSLDTTQTAEAQVETVIVTISGETGESPEC 1400
Db 1478 DAKKAIDAEAGKVTDADHDATLTAAQQAQQAQVADAEADKAKKAIDAAAGNADAIDQAKS 1537
Qy 1401 VGAHLLPAEKSSATGGHWTQLQHAEDTVPLGPESQAESTPIIIVTAPESTLHPDLOGELSA 1460
Db 1538 AGIKAIQDEQHKSGQ-----SIDTRKDDAKKAIDGVEAKITDAIDHDPTLT-----DAEKAT 1588
Qy 1461 SQRESSEEDKPDAGPDAGDKGKETAIEKVLKAEPEILELESKNK---IVLNVITQAVDQ 1517
Db 1589 QKQAVIAEADAKKAIDAAG--DADAVDQAQKAGIKAIQDQHKSGQALAIRKDAKKAIDE 1647
Qy 1518 PARTETAPETHAYD 1531
Db 1648 ----EAAKVSEAIID 1657
```

RESULT 10

```
US-09-815-242-5816
; Sequence 5816, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5816
; LENGTH: 2478
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5816
```

```
Query Match 4.7%; Score 377; DB 10; Length 2478;
Best Local Similarity 18.3%; Pred. No. 4.1e-09;
Matches 305; Conservative 274; Mismatches 625; Indels 462; Gaps 68;
```

```
Qy 47 PATKPKQKQLSVNGVAEQGDVHVEENQEQEE--EVVDVEDVQGESESDVRKD----101
Db 681 PTVKPAQKQDIIQAVTTRKQKQ---IKKSNASLQDEKDVANDKIGKIEKTKAIDDAATT 736
Qy 102 RVEEMANSTAVEDITKDGQGEETSEIIQIOPASENNVEEMVQPAESQANDVGFKKVPKV 161
Db 737 NAQVEAIKTKAINDINTTTPATTAK-----AAALEEFDEVVQAIDQA-----779
Qy 162 GFKFTVKKDKNEKSDTVQLLTVRKDEG--EGAEASVSGADHQEPSVETAVGESASSELK 220
Db 780 --PLNPPTTNEEVAEAIERINAAKVSQKAIETTTAQD-----LERVKNEEISKIENIT 832
Qy 221 QSTE-----KQEGTLKQESST-----EIPLOAESDQAABEAKDEG-----257
Db 833 DSITQKMDAYNEVKQAATARKAQATVSNATNEEV---AEAD-AAVDAAQKQGLHDIQV 888
Qy 258 EEKQEKETKSPSPSPVNSSETTSFKKFTTHGAGWRKKTSPFKSKEDDLETAERKE 317
Db 889 KSKQEVADTKS--KVLDKINAIQT-----QAKVPAADTEVENAYNTRK 930
Qy 318 QE-----AEKVDEEKEKTEPASEQEPAEDTDQARLSADYKVELPLEDQVQGLEASSE 372
Db 931 QETQNSNASTTEEQQAAYTDLTKKQEARTNLDAANTNSD-----970
Qy 373 EKCAPLATEVDEKMEAHQEVVAEVHVSVTEKTEEGGGGGAEGGVVVGSTGSLPPEK 432
Db 971 -----VTTAKDNSIAAINQVQAATTKKSKDAKAEIAQKA---SERKTAIEAMNDSTTEEQ 1021
Qy 433 LAEPQEVQAEPA-----BELMKSRMCVSGGDHTQLTD-----LSPKETLPKHEGI 482
Db 1022 QAAKDKVDQAVVTANADIDNAAANNDV-----DNAKTTNEATTAATPDANVPAKQAI 1076
Qy 483 VSEVEMLSQOERIKVQSGPLKFLSSGLKSLGKQKGRGGGDEEPGEYQHITESP 542
Db 1077 ADKVQ-----AGE-----TAIDGNNGSTTEEKAAKQOVQTEK- 1109
Qy 543 ESADEQKGESSASSPPEPETTTCLEKPLEAPQDGEABEGTTSDEKKREGITPWASF 602
```

```

Db 1110 TTADAALDAATNAEVAEAKKAALAK--IEAIQPAT-----TTKD-----NAKE 1151
QY 603 MWTPEKRVRPESDEKEELEKVKSAITLSSTDSVSEMODEVKTVGEBOKEEPRKRVDT 662
Db 1152 AIAATKAMERKATIAIQTQGITAEETIAANAVADVNATQANSINEANSQNDVQAKTGEN 1211
QY 663 SVSWEALICVSSSKKRAKASS-----DDEGPRILGDSHRAEASDKKGTG-- 713
Db 1212 SI--DQVLPYTKAKATANEETIAIINNKLOEIQATP-----DATDEEQADAEANTNG 1264
QY 714 ----AVPAST--OEODQAGSSSPBPAGSPBEGEVSPTWESPRLVTPR--KKSXKLE- 764
Db 1265 KANQIISATNTNAQVDEKAKANA-----EAINAVTKRVYKKAQAKBEI 1307
QY 765 EKADBSVEQLSTELEPESRESWSIKKIPGRKKRADGKQEQATVEDSGPVEINEDDP 824
Db 1308 DQLOATQTNVINNDQNAATTEKEAII-----QQLATA-----VTDAKN 1345
QY 825 NVPAVVPLENNAVREKEMAGNTELPQLGAVVSELSLTIVHTYSVAVIDGTBAVT 884
Db 1346 NITAA--TDDNGVDQADAGKNSIQSTQPATVAKSNKANDVDQAVTQONQAIIDTTGAT 1402
QY 885 SVEERSPSMISASVTEPLEHTAGEAMPVEEYTEKDIIAETPTVLTQTLPEKGD-AHDDM 943
Db 1403 TEE-----KNAKDLVLKAKKAYODILAAQT--TNDVTQIKDQAVADI 1444
QY 944 VTSEVDPTSEAVTATESEALRTEBTEASGAETTDVMSAV-----SQTLDSPDTEEA 998
Db 1445 QGITADTTIKVAVADELTATKANEGKALLAQATADATTEKEQONQVNDQLTGONINENA 1504
QY 999 TPVQEVESGVLDTEEEBQTOAI--LQAVADKVKESQVPATQVQTRGSKALEVEEVE 1056
Db 1505 QSIDDVNTA-----KDMAIQAIIDIQASTD-VKTNARAEELTEQO-----NKITEIL 1550
QY 1057 EDSEVLASEKEKEDVMPKGPVQAGAEHLAQSEGTQATPESLEVEBVTADVHVATCOVI 1116
Db 1551 NNNETTNEEKNDI--GPRVPAVEEG-----NNINATTTGDDVTATADTAQV 1596
QY 1117 KLQOLMEQAV-APESSETLTSETNGSTPLADSDTADGTQOD-----ETIDSQDSKATPA 1170
Db 1597 KVQOIHANPVKKPKAKKELDQAAADAKTKQI--EQTPNKSQEQEINDAKQEVDELINQAKTN 1654
QY 1171 VRQSCVTEEPATQAKKEPSTLPNNVPAOEHEGEEBGRDVLPTQOELTAAVPLVATE 1230
Db 1655 VDQSS-TNEYVDNAVKKEKAKI-NAVKTFSKXKD-----ALAKIE 1693
QY 1231 VGOESEVWMLGEKV-----KEBOEVFVHSG----- 1256
Db 1694 DAYNAKVNADNSNASTSEIAEAKQIAELKQIADQNVNATSKODILEVQIHNLDLIN 1753
QY 1257 ----PNSOKAA-----DVTYDSEVMGVAQOEKESTEVQSL----- 1288
Db 1754 DYTTPGKKEGATITDLYAVADQKKNISADTNATQDEKQQAQIKQVDQNVQALBESINNGV 1813
QY 1289 -----SLEBGEETDYEKEKRETKPE-----QVSEBGEQETAAPEHGTYKGVLLLD 1336
Db 1814 DNGVDVDDLTOGKAIDAIQVATATVKPRANQAIEVKAEIDTESIQSDQ----- 1862
QY 1337 MPSSRGALGSLGSPSLPDQDKAGCIEVQVQSLDTVTQTAFAVEK-----VIETV 1389
Db 1863 LTAERKTALAMI--KOITDAKQGI-----TDATTTAB-VERAKAQGLEAFDIT 1909
QY 1390 VISEGESPECVGAHLIPAEKSSAGCHWTLQHAEDTVPLG-----PESQABSI 1438
Db 1910 QIDST-----EKOKALE--ELEFALDQIEAGVNNVADATTEKEKFAFNAL 1952
QY 1439 PIIVTPEASTLHPDLOEISASQERESEEDKDPAGPDADKESSTAIEKVIAKAPELL- 1497
Db 1953 EDLISKATEDISDQTTNAEIAITVKNSALEQLKQRIINPEVKKNALAEIREVNVNKOIEIITK 2012
QY 1498 --ELESKNKIVLVNIQTAVNDQPARTETAPETHAIDSQO--VA 1538

```

```

Db 2013 NADADASAKETIARTDLGRYFDRPADKDKTQTNAEVABLQNVTI 2058
RESULT 11
US-09-815-242-12967
Sequence 12967, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haseibeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12967
LENGTH: 2478
TYPE: PR1
ORGANISM: Staphylococcus aureus
US-09-815-242-12967
Query Match 4.7%; Score 377; DB 10; Length 2478;
Beat Local Similarity 18.3%; Pred. No. 4.1e-09;
Matches 305; Conservative 274; Mismatches 625; Indels 462; Gaps 68;
QY 47 PATKLPOKNGQLSSVNGVAGQDVHVOENQGOEB-EVVEDVYGOSESEDEPREKD----- 101
Db 681 PTYVFOAKODIQAATVTRKQD-----IKSNASLDQDEKDVANDKIGKIEYKAIKIDIAATT 736
QY 102 RVEEMANSTAVEDITKQGOETSEIIEQIPASENNVEENVPAPAESQANDVGFKVFKFV 161
Db 737 NAQVEAIIKTAINDINQTTPATTAK-----AAALEEFDEVVQAQIDQ----- 779
QY 162 GFKPTVKKDKNKEKSDTYQLLVTKKDEG-EGABASVAGADHQBEPVEVRAVGSASKESEFLK 220
Db 780 --PLNPDTTNEEVAEAIERIRIAAVVSGVKAIEATTTADQ-----LERKNEIEIKENIT 832
QY 221 QSTE-----KOEGLTKOESQST-----ELIPQAESDOAAEBEAKXDEG----- 257
Db 833 DSTQKMDAYNEVQQAATARKAQANATVSNATNEEV--AEAD-AAVDAAQOGHLHDIOV 888
QY 258 EKOKEEPTKSPESPSSPVNSSETTSFKKFPTHGAWGRKKTSPFKSKEDDLETAEKKKE 317
Db 889 KSKQEVADTKS--KVLQKINAIQT-----QAKYKPADVEVENAAYNTRK 930
QY 318 QE-----AEKVDEEKEKTEBPASBQEPADTDQARLSADYKVELPLEDQVGLBASSE 372
Db 931 QEIQNSNSTTEEQQAATTELDTKQKEARTMLDAANTNSD----- 970
QY 373 EKCAPLAEVDEKMEHAEQEVAAEVHNVSTVEKTEBEGGSGGAEAGGVVEGTGSESLPREK 432

```


Db 971 -----VTTAKDINSIAAINOVQAATTKSDAKAEIAQKA-----SERKTAIEAMNDSTTEEQ 1021
Qy 433 LAEPQEVPOEABPA-----EELMKREMCVSGDHTQLTD-----LSPKTKLPKHPEGI 482
Db 1022 QAAKDKVDQAVVTANADIDNAAANDV-----DNAKTTNEATIAAITPDANVVKPAQAQAI 1076
Qy 483 VSEVEMLSQBERIKVQSGPLKLLFSSGLKLSGKKQKGGGGGDEBPGEYQHHTESP 542
Db 1077 ADKQV-----AQE-----TAIDGNGSTTEEKAQAQVQTEK- 1109
Qy 543 ESADQEGESSASSPEPEETTCLEKGPLEAPQDGEAEEGTTSQGEKKREGITPWASFKK 602
Db 1110 TTADAADAHTNAEVEAAKKAATAK--IEAIQPAT-----TTKD-----NAKE 1151
Qy 603 MVTPEKVRPPSSDKKEELEKVKASLUSSTDSTVSEMQDEVKTVGEEQKPEEPKRRVDT 662
Db 1152 AIATKANERKTAIAQTQDITAEIEAANADVDNAVTAQNSIEAANSQNDVQAKTTGEN 1211
Qy 663 SVSWEALICVSSKKRARKASS-----DDEGPRTLGDSHRAEASKDKEAGTD-- 713
Db 1212 SI---DQVTPVTKATARNETAILNKKLQEIQATP-----DATBEKQAADEANTENG 1264
Qy 714 ----AVPAST--QPQDQAQSSSPPEAGSPGEGSVTWESFKRLVTPR--KKSLSKLE- 764
Db 1265 KANQAISSAATNAQVDEAKANA-----EAINAVTPKVVKQAQKDEI 1307
Qy 765 EKABDSSVEQLSTIEIPRESBWSVSKFIPGRKKRADGQEQATVEDSGFVINEBDDP 824
Db 1308 DQLOATQTNVINNDQATTEKEAAI-----QQLATA-----VTDKKN 1345
Qy 825 NPAVAVPLSEVNAVEREKMEAGQNTLPGALLGAVVSEELSKTLVHTVSVAVIDGTAVT 884
Db 1346 NITAA-----TDDNGVDQAKDAGKNSIQSTQPATAVKSNKNDVQAVTQNOAIDNTTGT 1402
Qy 885 SVEERSPSWISASVTEPLEHTAGAMPVPERVEBTKDIIAETPTVLTLQTPBKGD-AHDDM 943
Db 1403 TEE-----KNAAKDLVKAKKAYQDILNAGT---TNDVTQIKDQAVADI 1444
Qy 944 VTSEVDTSEAVTATSEALRTEVEVTEASGAETTDMSAV-----SOLDSPDPTTEEA 998
Db 1445 QGITADTTIKDVAKDELATKANEQKALIAQTADATTEKEQANQVDAQLTQGNQNIENA 1504
Qy 999 TPQVEVSGVLDTTEERQTOAI--LQAVADKVKKEEQVPAQTQVORTGSKALEKVEVE 1056
Db 1505 QSIDDVNNTA-----KONAIQAIQDIOASTD-VKTNARAELLTMOQ-----NKITEIL 1550
Qy 1057 EDSEVLASEKEKVMKPGPVOEAGAEHLAQSGSETQATPESLEVPEVTADVHVATCOVI 1116
Db 1551 NNNETTNEEKNDI--GPVRAAYEEGL-----NNINAATTGVDVTTAKDTAVQ 1596
Qy 1117 KLOQLMEQAV-APESSETLTDSETNGSTPLADSDTADGTQDQ-----ETIDSDQSKATAA 1170
Db 1597 KVQQLHANPVKKPAGKXELDQAAADKKTQI--EQTNPASQOEINDAKQEVDTTELNAQKTN 1654
Qy 1171 VRQSVTEEEAATAKEEPSTLPNNVPAQEBHGEPPGRDVLPTQQLTAAAVPVLAKTE 1230
Db 1655 VDQSS--TNEYVDNAVKEGKAKI-NAVTFSEYKKD-----ALAKIE 1693
Qy 1231 VQBEVGDWLDGKGV-----KEQEVFVHSG----- 1256
Db 1694 DAYNAKVNADNSNASTSSEIAEAKQKLAELKQTDQNVQNAQTSKDDIEVQIHNLDLONIN 1753
Qy 1257 ----PNSQKAA-----DVTYDSEVMGVAGCQBEKSTEVQSL----- 1288
Db 1754 DYTPTCKESATDLVAYADQKKNNISADTNATQDEKQQAQKQVDQNVQTALESINNGV 1813
Qy 1289 -----SLBEGEMETDVEKEKRETKPE-----QVSEBEGEQTAAPEHGTGYPVLTLTD 1336
Db 1814 DNGVDVDDALTQGKAADIAIQVDAVTKPKANQAIQVKAEDTKESIDQSDQ----- 1862
Qy 1337 MPSSERKALGSLGGSFSLPDQDQAGCTEVQVQSLDVTVTQTAEAVEK-----VLETV 1389
Db 1863 LTAEKTEALAMI---KQITDOAKQGI-----TDATTTAB-VEKAKAQGLEAFDNI 1909

Qy 1390 VISETGSPSCVGAHLLPAEKSSATGHHWTLOHAEDTVPLG-----PESQAESI 1438
Db 1910 QIDST-----EKQKAIE---ELETALDQIEAGVNVNADATTEEKEAFTNAL 1952
Qy 1439 PIIVTPAPESTLHPDLOGEISASQRESSEEDKPDAGDADGKRESTAIEKVLKAEPEIL- 1497
Db 1953 EDILSKATEDISQTTNAEIATVKNSALEQKKAQRIINPEVKKNALAIREVNVNKOIELIK 2012
Qy 1498 --ELESKNKIVLNVIOITAVDQFARTETAPETHAYDSQTQ---VPA 1538
Db 2013 NADADASAKAIEARTDGLGRYDFRADKLDKTKTQNAEVAELQNVITIPA 2058
RESULT 12
US-09-919-172-98
; Sequence 98, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 98
; LENGTH: 3256
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 2700132CD1
US-09-919-172-98

Query Match 4.7%; Score 377; DB 10; Length 3256;
Best Local Similarity 20.2%; Pred. No. 5.6e-09;
Matches 387; Conservative 241; Mismatches 762; Indels 524; Gaps 83;

Qy 14 QPAGSDTPSELVLVLSGH-----GPAABASGAAGDPADADPATKLPQKNGQL 58
Db 1412 QTSGETTHTDKVPGGDKSINAFRETAQKLDPAASVTSGRHP-----KTKKAQPL 1464
Qy 59 SSVNGVAEQGDVHVQENQEQEE-----EVVDEDVQGR--ESEDVREKDRVEEM 106
Db 1465 EDLAGWKELFQTPVCTDKPTTHETKTKIACRSQPDVPTTSSKPKQSRSLURKVDVBEF 1524
Qy 107 AANSTAVEDITK-----DQOE-----ETSEIIEQIPASENNVEENVQPAES 147
Db 1525 FALRKRTPSAGKAMHTPKPAVSGEKNIYAFMGTPVQKLDLTENLTGSKRRLQTPKEKAQA 1584
Qy 148 QANDVGKVKFVGFKFTVKDKNEKSDTVQLLTVKKDEGEAEAS-----VGA 197
Db 1585 LEDLAGFKELFQTRG--HTESMTNDKTAQVACKSSQPDLDKNPASSKRLKTSGLKGVG 1642
Qy 198 GDHQEPSVEVAVGESASKESELKQSTEQKGTLEKQESSTEIPLQAESDQAAEEAKDEG 257
Db 1643 KEEL-----LAVGKLQTSGETTHTHTEPTGDKGSKMAFMESP-----KQILDSAAALTG 1692
Qy 258 EEQKEKEPTKSPSPSPVNV---SETTSSFKKFFTHGWAGWRKTKTSFKKSKEKDDLETAE 313
Db 1693 SKQLRTPKGKSEVPEDLAGFIELFQTPSHTKESMTNEKT---TKVSYRASQPLVDTPPT 1749
Qy 314 KRKEQ--EAEKVDEEEK---EKTEPAS-----EBOE-----PAEDTDQ- 346
Db 1750 SSKPQPKSRSLRKADTEEEFLAFRKQTPSAGKAMHTPKPAVGEKNDINTFLGTPVQKLDQP 1809
Qy 347 -----ARLSADYEKVELPLEDOVGDL-----ASSEK-----CAPLATEVFD 384
Db 1810 GNLFSGNRRLOTRKEKAQ-ALEELTGPRELFTQPTCTDNPTTDEKTKKILCKSPQSDPAD 1868

```

Qy 385 -----EKMENQEVVA--EVHVSVEKTEEEQGGGAEAGVVEGTGSLP 429
Db 1869 TPINTKOPKRLKKAADVEEBFLFRKLTJPSAGKAMHTPKAAGEKDINTFVGJ----P 1924
Qy 430 PEKJ-----AEPOEVOEAPAEELMKSRMVCVSGDHQ-----LTFDS--- 469
Db 1925 VEKJDLGNLPGSKRRPOTPEKAKALEDLAGFKLFOTPG--HTEBSMTDDKITEVSKS 1983
Qy 470 --PEEKTLPKHREGIVSEVEMLSQERIKVQGSPL--KLFSSSGKLKSGKKQKRG 524
Db 1984 PQPBPVKP-----TSSKQRLKLSLKGIVGKEEVLVPGKLTQTSGETTQTHRE 2031
Qy 525 GGGDEPEEYOHITESPESA--DEQKESASSPREP--EETTCLE-----KGPLAEAD 576
Db 2032 TAGD-----GKSIAKAFKESAKQMLDPANVGTGMEWRPRTPEEKQSLEDLAGELEFOTPD- 2087
Qy 577 GEABEGTTSDEGKKREGJTPMASFKQWTPPKRVRP-----SESKEELEKVKSATLSS 632
Db 2088 -HTEBSTDDTKTKIACKSP--PPESMDTPSTRRRPKTPLGKRDIIVELSLAKQLTQTT 2144
Qy 633 --TD-----STVSEMODEVKTV--GEQKPEEPKRRVDTSVSEAL----- 669
Db 2145 HTDKVPDEDKGINVFRFAKQKLDPAASVTSKQOPRTPKGAKQPLEDLAKELFOTP 2204
Qy 670 ICVG-----SSKKARAKASSDDGCGPTLGGDSHRA-----EASKDKAGTDAV-- 715
Db 2205 ICTDKPTTHEKTKIACKSPQPPDPVGTPTIFPKQKRSRKADVEEESLALRKRTPSVGK 2264
Qy 716 -----PASTOEOPOAQSSSP--EPAGSPSEEGVSTWESFKRLVTPRKRSKSKLEKAE 768
Db 2265 AMDTPKPPGCGDEKMAKMGTFVQKLDLPGNLPGSKRMPQ-----TPKKAQA--LEDLAG 2318
Qy 769 DSSVEQLSTELPSREESWVSJ-----KKFIPGRKKRQADGQEQATVED 813
Db 2319 FKELFQGTGKTPTDEKTKIACKSPQPPDVDTASTKQPKRMRLRADVEEELALRK 2378
Qy 814 SGPEVINEDDPNPAVNPPLSEYNA--VER--EKMENQGN-----TELPGLLGAVY 859
Db 2379 RTPSAGKAMDTPKPAVSDKMNINTFVEETPVQKLDLGNLPGSKROPQTPPEKAEALIEDLV 2438
Qy 860 VSEELSKTLVHTVSAVAVIDGTRAVTVSVEERSPSWISASVTEBLEHTAGAMPVEEVTEK 919
Db 2439 GFKELFOTPGHTEBSMTD--KITEVSKSP-----QESKRTSSSKQRLKIPLV 2487
Qy 920 DIIAETTV-----LTQTLPEGKAHD-----MVTSEVDTSEAVTATETSEALR 965
Db 2488 KYDMKEEPLAVSKLRTSGETTQTHTEPTGDSKSIKAFKESPKQILDPAASVTSGRQRLR 2547
Qy 966 T-----EYVTEASGAE--TTDMSAVSOLTUSPTTEATVQEVESGV 1008
Db 2548 TRKEKARALEDLVDPEKELFSAGHTEBSMTIDKNKTIPEKSPPELITDPAATSKCPKTR 2607
Qy 1009 LQTEEEBQTOAILLAVADKVEESQVPAQTVQRTGSKALKVEEVEDSEVLASEKEX 1068
Db 2608 LRKEKVEE--LSAVERLTQTSOSTHKEPPASGDEGIVLKORAKKKNPVEEBS 2662
Qy 1069 DVMKGPVQEA--GAEHLA--QGSSETQATPES-----LEVEVTAADVDA 1109
Db 2663 RRRPAPKPKAQLPLEDLAGFTELSETSGHTQSLAGKATKIPCESPLEVDDVTASTR 2722
Qy 1110 VATCOVILKQOLMEQAVAPESSETLTDSETNGSTPLADSDTA----- 1151
Db 2723 HURTVQKQV-----VKEEPSAVKFTQTSGETTDADKEPAGEDKGIKALKESAKQTPA 2775
Qy 1152 -----DGTQOQDETI-----DSQDSKATAVROSOVTEEBA 1182
Db 2776 PAASVTGRRRRPRAPRESAQALIEDLAGFKDPAAGHTEBSMTDKTKIPCKSPLELDTA 2835
Qy 1183 TAOKEEPTLPPNNVPAQEEHGEEPGRDVLPTQOELTAAAVPVLAKTEVGOGEVDMLDG 1242
Db 2836 TSSKRRPTRAKQVKEE-----LLAVKLTQTT-----SGETTHMDX 2873

```

```

Qy 1243 EKYKEOEVFVHSGPNSOKADAVTYDSE--VMG-----VAGCOEKESTE 1284
Db 2874 EPVEGKGTAKFKQPARK-----LDABDVIGSRORAPRPAKQAPLEDLASFELSOTP 2928
Qy 1285 VQSLSEEGEME--TDVEKEKET--KPEOVSEEGEOETAAPHEHGTGKPYLTLDMPSE 1341
Db 2929 GHTLELANGADSPSAPKQPPDSGKPLKIS--RVLRAPKVE--PVGDVYSTRD--PVKS 2983
Qy 1342 RGRALGSLGSPSLPDDQKACIEVQOSLDTTYQTAAE-----AVEKVIETVVISERGE 1396
Db 2984 QSKNTSL--PPLPFRRGG-----KQSVTGTKRLCWPAPBEIYEE----- 3024
Qy 1397 SPECAGHILPRAKSSATYGHMTLOHAEDEVPLQAPESQAESIPITVPAPSTLAPDLOG 1456
Db 3025 -----LPASIK-----QRAVPRARAGSSSEPVVMKSLRISAKRIEPAE 3063
Qy 1457 EISASQRESEEDK--PDAGDADG--KESTALIEKYLAEPEILELESKNKIVANVQTA 1514
Db 3064 ELNSNDMKTNEKHLDQSVENKGISLRSRQKTEABEQOITEVFVLAERIEIN----- 3118
Qy 1515 VDQFARTE-----TAPETHAV--DSQTPACRLDSREPNRCWTRMDAKKHP 1561
Db 3119 -----RNEKKEMKTSPEMDIQNPDDGARKPIPRDKVTENKRLCSARQNESSQP 3167

```

```

RESULT 13
US-09-764-176-7
; Sequence 7, Application US/09764176
; Patent No. US20020127553A1
; GENERAL INFORMATION:
; APPLICANT: NOTEBORN, Mathieu Hubertus Maria
; APPLICANT: DANEN-VAN OORSCHOT, Astrid Adriana Anna Maria
; APPLICANT: ROHN, Jennifer Leigh
; APPLICANT: WEISS, Bertram
; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN
; FILE REFERENCE: 4725US
; CURRENT APPLICATION NUMBER: US/09/764,176
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1400
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Description of Sequence: Amino acid sequence deduced from the nuc
; OTHER INFORMATION: Ictc acid sequence of AAP-
US-09-764-176-7

```

```

Query Match 4.6%; Score 368; DB 10; Length 1400;
Best Local Similarity 19.6%; Pred. No. 5.2e-09;
Matches 283; Conservative 226; Mismatches 495; Indels 442; Gaps 65;

Qy 95 EDVREKDVEMAANSTAVEDI--TKDQGEETSEIIEQIPASENNVEWQAPAESQANDV 152
Db 92 DNLKFNKLTINEEDADTMRLQPIGRDKGLMTWYQLDQ-----DHVNRVYIEQDDQDS- 145
Qy 153 GFKKVFEVGRFTVKDKDNKSDPTVOLLTVK-----KDEGEAGASVAGADHOEPSVE 206
Db 146 -----SWKCIY--RNRNELAETLALLKAQIDPVLKNSQQONS-----RSPLEL 190
Qy 207 TAVGSASKESELKOSTEKQCTLKQEOSSSTEIPLQAESDQAAE--EAKDGEKQKE 264
Db 191 --DETKKEERTPKQEOQKE--SEKKKSEEOPMWDLNENRSTANVLEETTVKKEKEDKE 244
Qy 265 PTK-----SPSSSPVNSSETTSSFKKFFHGAAGRKKTTSFKSKEDDLET--- 311
Db 245 LVKLPLVILKEKLPENNEKKTIKEESDSFKENYK-----PIKVEVKECRADPKOTSS 298
Qy 312 AEKKEQEAQVD-----EEKEKTEPASEQEPALDTQD-----RLSADYE--- 354
Db 299 MEKPAQEPERIEFGNKKSHHEITKSTETETKLNQDQAKIPLKKEIKLSDDFDPSV 358

```

QY 355 -----KVLPLEDOVGLERASSBEKCAPLAT-----EVFDEKMEAH----- 390
Db 359 KGPLCKSVTPTKFELDKIQRKEETCKRIISTITGALGHEGKQLVNGEVSDRVAFPNFKTEP 418
QY 391 -----QVWAEHVSTVEKTEEGGGEAGBGVVVE-GTG-----ESLPPK----- 432
Db 419 IETKFYETKEESYSPKORNIITEGNTGESLNSVITSMKTGELKETAPURKODADSSISV 478
QY 433 -----LAEOVEPQE-----AEPAEEL-----MKSREMCVSGGDH----- 462
Db 479 LEIHSOKAQIEEPPPEMETSILDSSEMAKOLSSKTALSSSTESCTMKGEKSPKTKDKRPP 538
QY 463 -----TQLTDLSPKTLPK-----HPE 480
Db 539 PILECLEKLEKSKTFLDKAQLRSLPIPEEVPKSTLSEKSPGSEAAETSPPSNIIDHCE 598
QY 481 GIYSEVEMLSQERIKVQGSPLKLFSSGGLKLSGKKQKGGGGGDEBPGEVQHIHTE 540
Db 599 KLASEKEVVCQSTSTVGGQSVKKV-----DUETL-----KEDSEFTKVEMD 640
QY 541 SPESADQKGESSASSPEEPTTCLBKPLE-----APQDGEABEGTTSDGKKR 591
Db 641 NLDNA-----QTSGIEEPESET-----KGSQKSKFYKLVPEEETTASENTEITSERQK 689
QY 592 EGI-----TPWASFCKMWTPKRVRRP-----SESQKEBELEKV-----KSATLSSTDSTVSE 639
Db 690 EGIKLTIIRISRRKKPDPSPKVPLEPENKQBEKEEKTNVGRTLRRSPRISPTAKVAE 749
QY 640 MODE-----VKTVGEEOKPEE-----PKRVDTS-----VSWEALIC 671
Db 750 IRDQKADKKGEGEBEVEESTALQTKDKKILKSEKDTNSKVSKPKGKVRWT----- 805
QY 672 VGSKKRARKASSDDGGGPRTLGGDSHRAEASKEAGTDAVPASTOBODAOAGSSSP 731
Db 806 --GSRTEGRWKYSNDE-----SEGSSEKSSAASEEE-----EKSEEAAILADDD 850
QY 732 EP---AGSPGEGVSTWSEFKR-----LVTP-----RKSKSKLEKA 767
Db 851 EPCKKCLPNHPILLICDSGYHTACLRPPMLIIPDGEMFPCPCQHQHLLCKLEEQ 910
QY 768 EDSSVEQLSTELPSPRESVW---SIKKETP-----GRKKRADGQEQATVEDSG 815
Db 911 QDLVDALKKERARRKRLVYVIGISLENIPQEPDFSEDBEKKDKSKKSKANLERR 970
QY 816 PV-----EINEDDPNPAVPLSEYNAREKMEAQNTLPLLGLAVVYSELSKT 867
Db 971 STRTRKICSVRFDEFDAID-----EAIEDDIKEADG-----GGVGRGKIDISTI 1014
QY 868 LVH---TVSVAVIDGTRAVTSVEERSPSWISASVTEP-----LEHTAGEAMPP 912
Db 1015 TGHGKDIS-TILDEERK-----ENKRPQRAAAARRKKRRRLNLDSDNLDEESEDEFK 1069
QY 913 VEEVTEKO-LIAEETPVLTQLEPGKDAHDMVTSEVDFTSEAVTATETSEALRTBEVTE 971
Db 1070 ISGSDQEFVVDENPDESEDDPPSND-----DSOTDFCSRRLRRHPSRPMQSRRLR 1123
QY 972 ASGAETTDMSVAVSOLTDSPDTTEATPVQEVESGVLDT-----EBEERQTAIILQAVA 1026
Db 1124 KTPKKYSD-----DDEESEEENS---RDSSEDFSDDFDVFETRRRRRRNRKQK 1173
QY 1027 DKVKESSQVPATQTVORTGSKALEKVEEVEDSEVLASEKEDVMKPGPV-QBAGAEHLA 1085
Db 1174 INYKEDSESDGSKSLRRG-KEIRRVHK-----RRLSSSESESYLSKNSDEDELAKESKR 1228
QY 1086 QGSETGQATPESLEVPVADVDHVTACQVILQQLMEQAVAPESSETLTDSGTNGSTPL 1145
Db 1229 SVKRGKSTDEYSEADDEEEEBEG-----KPSRKRLHRIETDEBESCDNAHGDAQPA 1281
QY 1146 ADSDTADGTQOQDITSDSKATAAVRQSQVTEEEAATAQK-----EPSPSTLPNN 1195
Db 1282 RDS-----QPRVLFSEQUESTKKPRIESDEBEDFENVGVGSPLDYSLVDLPST----- 1330

QY 1196 VPAQEBHGEBPGRDVL-----EPTQQ-----ELTAAAVPVLAKEVGOBEVDMLDGEKV 1245
Db 1331 -----NGSPGKRAIENLIGKPTKQTPKDNSTAGSLASNGTSGQEA-----GAPE 1378
QY 1246 KEQOE 1251
Db 1379 EEEDEL 1384
RESULT 14
US-09-963-875-1
; Sequence 1, Application US/09963875
; Patent No. US20020164307A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Stem Cells of the Islets of Langerhans and Their Use in Treating
; TITLE OF INVENTION: Mellitus
; FILE REFERENCE: 17633/1235
; CURRENT APPLICATION NUMBER: US/09/963,875
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US60/169082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/215109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/238880
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/731261
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-875-1
Query Match 4.4%; Score 354; DB 9; Length 1618;
Best Local Similarity 20.1%; Pred. No. 2.5e-08;
Matches 342; Conservative 224; Mismatches 623; Indels 510; Gaps 75;
QY 183 VKKDEGEAGASVAGDHPSPVETNAVGESASKE-----SELKQS--T 223
Db 26 VKALEEQNELLGAGLGRQSRQADTWSRAHADDELAALRALVDQRWREKHAAEVARDNLA 85
QY 224 EKOEGT-----LKQEOSSTEI-----PQAE-----SDQAAEEBAKDS----- 256
Db 86 EELEGVAGCEQLRLARERTTEEVARNRAVRAEAKCARAWLSQGAELERELEALRAVHE 145
QY 257 ----GEEKQEKFTKSPSPSPVNSSETTSFKKFFTHGWAG-----WRKKTSPFK 303
Db 146 EERVGLNQAACAPRLPAPRPAPAPAEVEELARRLGEAWRGAVRGYQERVVAHMETSLDQ 205
QY 304 SKEDDLTAEKKE--QEAKEVDEEKEKTEPASEBQEPADTDQARLSADYKVKVLPLE 361
Db 206 TRERLARAVOGAREVRLELQQLQAERGGLLERRAALQORLEGRWQERLRAT-EKFQAVE 264
QY 362 DQVGDLASSEKCAPLATEVDFDEKMEAHQEVVAEVHVST-----VEKTEBEEQGGG-- 413
Db 265 ----ALEQEKQGLQSOIAQVLEGRQOLAHKWSLSLEVAITYRTLLEAENSLQTPGGGSK 320
QY 414 -----EAGGVVVEG--TGESLP---PEKLAEPQEVQPEAB-PAEELMKSRMVCV 457
Db 321 TSLSFQDPKLELQFPRTPEGRRLGSLLPVLSPTSL--PSPLATLETVPYPAFLKNGEFLQ 378
QY 458 SGGDHQTLDLSPEKTLUPKHPEGIVSEVEMLSQERIKVQGSPLKLFSSSGLKLSK 517
Db 379 ARTPTLASTPIPP---TPQAPSPAV-----DAEIRAQDAPLSLLQTOGGRQKAPER 426
QY 518 KQKGRGG-----GGDEEP-GEYQHIHT-ESPESADEQKGESSASSPEPETTCLEK 569
Db 427 LRAREARVAIPASVLPGEPPGQEQEASTGQSPED-----HASLAPPLSPDHS----- 475
QY 570 PLEAPQDGEABEGTT-----SDGEKKREGITP-----WASFKMVTPKKRVRRPSE 615

Db 476 -LEA-KOGESEGSRVFSTICRBEQOIGWGLVEKTAIEGKVSVSSIQOIMEEDLNKEI 533
Qy 616 SDKEEELKXVSATLSTDSYVSEMDEKTVGE-----EQKEPEPRKRVDTSVSMAL 669
Db 534 QDSQVPLEK-----ETLKSIG-----EIIQESLKTLENQSHETLERENQECRPSLEEDL--ETL 585
Qy 670 ICVGSSKKRARKASSDDEGGR-----TLGDSHAEFASAKDKEA----- 710
Db 586 KSLKKNKRAIKGCGSETSRKRGROLKPTGKEDTQTLQIKENQELMKSLEGNLETF 645
Qy 711 ---GTD--AVASQTOEQDQ---AOGSSSEPAQSPSEBGVS----- 744
Db 646 LPPGTNOELVSSLOENIESLTALKEKNOEPLRSFEVGDDEKRLPLTKENQEPILASLEDE 705
Qy 745 TWSEFKRLVTPRKSKSKLKEKADSVYQULSTE-----LEPSEBSWSIKKIFIGR 797
Db 706 NKEAFRLSEKNOEPLKTLBEE--DQSIVRPLETENHKSLSLEBODQETLRLER--ETQ 762
Qy 798 RKRKADGQEOATVBDSGFVEINEDDPNVPAVPLSEINAVREKMEAGQNTLPOLLGA 857
Db 763 QRRSLGQDQWTLRP-----PEKYDLEPLKSLDOE----- 793
Qy 868 VYVSEELKTLVHTVSVAVIDGTRAVTSVEERSPSMISAVTEPLEHTAGAMPVEEVT 917
Db 794 -----IARPLENNOEPLKSIKES 813
Qy 918 EKDIETPVLVTLPEKGAHDMVTSEVDTSEAVATETSEALTEETVTEASGAE 977
Db 814 VEAVSLETEIIESLSKAGQ-----NLETLSPTQOPLWPEELINSGNE 861
Qy 978 TTDWVSAVSQULDSPTTEATPVOEVESGVID-----TEEE 1014
Db 862 SSRKNSTRITGVCGSEPRDIQTPRG-ESGIIIEISGMEPGEFISRGVDKESORNLBEE 920
Qy 1015 EEOQTALIQAVADKXKEEQ--VPAQTVOYRTGSKALEXEVEVEBDEV-LASEKEXVMP 1072
Db 921 ENVGKGEYQESRSLEEBQOELPQSHADVOR--WEDYKQOELAOESPPMGAVENKD--- 976
Qy 1073 KQPVQACAEHLAOGSETQOATPESLVEPVTADVD-----HVATCQVYIKQOIMEQA 1125
Db 977 -----EAEINLEBOQGFTR--EEVVEQELNATEEVPFGEGHPENPKEQGLVEGA 1029
Qy 1126 VAPESSETLTDSETNGS-----TPLADSTPADGICQ--DETIDSQDS 1165
Db 1030 SVKGAEGIQDBEGSQOQVTPGLQAPQGLPEALIBLVDDVAFGDDQASPEVMLGSEBA 1089
Qy 1166 KATAAV-----ROSQVTEEAATQKEPSTLPNNVPAOE----- 1200
Db 1090 MGESAGAGPGLGQGVGLGDPGHLTREVMEPPLBESLBAKRVGLGPKKDLBEEAG 1149
Qy 1201 ---EHGEERG--RDVLEPTQO---ELTA---AAVPLAKTEVQEGEVMDLGEKV 1245
Db 1150 LGTERSELFGKSRDPMERPREGRESEAPAPRAEBAFAETLGHGSAAPSWPLGSE 1209
Qy 1246 KEEQ-----EYFVHSGPNSQKADVITYDSEVNGVAGCCQKESTEVQSLBEG 1293
Db 1210 AEDDVPPLVSPSPYTYTLPLEDAPGLQPOAESQFAS--WGVQRAEAGKYESQOELGSG 1268
Qy 1294 EMETDVEKEKREKRPQVBEQGEQTAEPHE--GTYGK-----PVLITDMPSEBG 1343
Db 1269 EIIPEGIQEGEESRER--SBEDELGETLPDSTPLGYLSSPSPMTPLESRGHPKETE 1326
Qy 1344 K-----ALGSLG--GSPSLPDODKAG--C-----LEVQVOSLDTTYT--OTAEA 1381
Db 1327 KEGWDPAYLALSEGLEPSEKEEGBEGEGGRSDLSSEEDLGTAPLPGVGAVER 1386
Qy 1382 VEKVLTEVV-----ISETGESE-----CVGAHLLE 1407
Db 1387 LQGVPLLLDPAAMDRDGEDGPADEBESGEGEEDQEEGRFEGAGRWGGSVGS--LQ 1444
Qy 1408 AAKSSATGHWITQHAEDYVPLGPESQABEIPITVYAPASTLHAPLQ--GELISAQORER 1465

Db 1445 ALSSSQGEFLESDSVSVPMW-----DSLKGAVAGAPKTALETESQDASBPSSGEES 1499
Qy 1466 S-----EEDK-----PDAGPDADKESFALIKVLAPELLELSKSKXI---V 1507
Db 1500 DVLSEBEDKPGPLEIPSGMEDAGPAD-----IIGVNGQGNLKGSKSHVANGV 1550
Qy 1508 LNVIGTAVDQFARTETAPE 1526
Db 1551 MNGLEQSESGARNALYSE 1569

RESULT 15
US-09-864-761-36182
: Sequence 36182, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aeomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/160,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 36182
: LENGTH: 617
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC005529.7
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.94
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:06:25 ; Search time 59 Seconds
(without alignments)
2600.518 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSSTQSRPEQAGSDT.....AWAQRKCLRLQLKAPVSK 1596
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	5421.5	67.2	1346	2 A57376	probable regulator
2	4083.5	50.6	1684	2 JW0057	gravin - human
3	585.5	7.3	1829	2 T24583	hypothetical prote
4	574	7.1	5327	2 T13564	microtubule-associ
5	501	6.2	2464	1 QRMSP1	microtubule-associ
6	497.5	6.2	3488	2 T34418	hypothetical prote
7	493	6.1	3924	2 S37431	ankyrin 2, neurona
8	489	6.1	6642	2 T29757	protein UNC-89 - C
9	488	6.0	5170	2 T15348	hypothetical prote
10	485	6.0	2484	2 T26216	hypothetical prote
11	485	6.0	2607	2 T26215	gene 11-1 protein
12	483.5	6.0	1948	2 S00485	microtubule-associ
13	470	5.8	2364	2 A56577	probable heat choc
14	466.5	5.8	1871	2 D96796	nestin - golden ha
15	456.5	5.7	1804	2 T34518	R27-2 protein - tr
16	438	5.4	1128	2 T30296	mature-parasite-in
17	434	5.4	1526	2 A45605	microtubule-associ
18	433.5	5.4	2774	2 A43359	NF-180 - sea lamp
19	425	5.3	1110	2 T51116	364K Golgi complex
20	419	5.2	3187	2 JC5837	chondroitin sulfat
21	411.5	5.1	3582	2 A47171	plectin - rat
22	408.5	5.1	4687	1 A39638	myosin heavy chain
23	407.5	5.0	2139	2 T18296	liver stage antige
24	403.5	5.0	1909	2 A45592	elastic titin - hu
25	402.5	5.0	7962	2 I38446	hypothetical prote
26	402	5.0	2218	2 B84583	plectin - human
27	401.5	5.0	4574	2 G02520	plectin [imported]
28	400.5	5.0	4684	2 A59404	microtubule-associ
29	400	5.0	1830	2 A37981	

30	400	5.0	3225	2 I52300	giantin - human
31	397	4.9	3507	2 T34513	hypothetical prote
32	395.5	4.9	1825	2 T31507	microtubule-associ
33	393	4.9	3259	1 A56539	giantin - human
34	392.5	4.9	1822	2 S33441	EF protein - Strept
35	389	4.8	1262	2 T25233	hypothetical prote
36	389	4.8	1621	2 A82255	hypothetical prote
37	388	4.8	1805	2 A34736	nestin - rat
38	385.5	4.8	1087	1 QPM5H	neurofilament trip
39	382.5	4.7	1320	2 JC5630	TCOP1 protein - mo
40	382	4.7	2116	2 A26655	myosin heavy chain
41	380	4.7	1558	2 B71603	RESA-H3 antigen pF
42	379	4.7	2094	2 S33124	tpr protein - huma
43	379	4.7	2722	2 T20532	hypothetical prote
44	379	4.7	2738	2 E88320	protein F07A11.6 l
45	377	4.7	1634	2 T26517	hypothetical prote

ALIGNMENTS

RESULT 1

A57376
Probable regulatory protein 322 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 08-Feb-1996
C:Accession: A57376
R:Lin. X.: Nelson, P.J.; Frankfort, B.; Tomblar, E.; Johnson, R.; Gelman, I.H.
Mol. Cell. Biol. 15, 2754-2762, 1995
A:Title: Isolation and characterization of a novel mitogenic regulatory gene, 322, which encodes a protein with conceptual translation
A:Reference number: A57376; MUID:95257957; PMID:7739556
A:Accession: A57376
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1346 <LIN>
A:Cross-references: GB:U23146
C:Genetics:
A:Gene: 322

Query Match 67.2%; Score 5421.5; DB 2; Length 1346;					
Best Local Similarity 91.6%; Pred. No 5.5e-187;					
Matches 1098; Conservative 24; Mismatches 62; Indels 15; Gaps 7;					
QY	387	MEAHQEVAVHVS	VTVEKTEEEQGGGEA	GGVVGVTGTS	LPPEKLAEPQVQBAEPA 446
Db	1	MEAHQEVAVHVS	VTVEKTEEEQGGGEA	GGVVGVTGTS	LPPEKLAEPQVQBAEPA 60
QY	447	EELMKREMCVSG	GDHTQLTDLSP	EKTLPKHPEGIV	SEVEMLSQERIKVQSP
Db	61	EELMKREMCVSG	GDHTQLTDLSP	EKTLPKHPEGIV	SEVEMLSQERIKVQSP
QY	507	SSSGLKLSGKK	QKRGKGDEE	PGCYQHHTSP	ESADEQKGESSASSP
Db	121	SSSGLKLSGKK	QKRGKGDEE	PGCYQHHTSP	ESADEQKGESSASSP
QY	567	EKGLPEAPQD	GEAEETSDG	EKKRE---	GITPWSAFKQWTPK
Db	181	EKGLPEAPQD	GEAEETSDG	EKKRE---	GITPWSAFKQWTPK
QY	623	EKVKATLS	TDSTVSEMQ	DEVTVEGEQ	PEPKRVDTSVSEW
Db	235	EKVKATLS	TDSTVSEMQ	DEVTVEGEQ	PEPKRVDTSVSEW
QY	683	SSSDDEGP	PTLGGD	SHRAEAS	SKDEAGTDVAPAST
Db	295	SSSDIR	-GPTLGG	QSGQSR	QRRSRDTPAST
QY	743	VSTWSEFK	RLVTRPKSK	SKLEEK-	AEDSSVEQLSTIE
Db	354	VSTWSEFK	RLVTRPKSK	SKLEEK-	AEDSSVEQLSTIE
QY	800	KRADGQ	EQATVSD	SGPEIN	DDPNVAVPLSEY
Db					

Db 414 KGQMGROCATVEDSGPVEINDEEDVPAVNVPLSEYDAVERKMEAGNAELPSCMGCV- 472
 Qy 860 VSEELSKTLVHTVSVAVI DGT RAVTSVEERSPSWI SASYTEBLEHTAGAMPVEEVTEK 919
 Db 473 VSEELSKTLVHTVSVAVI DGT RAVTSVEERSPSWI SASYTEBLEHTAGAMPVEEVTEK 532
 Qy 920 DIIAETPVLTOTLPEGKAHDMDMTSEVDFPFSSEAVTATETSEALRTEEVTAASGAEEET 979
 Db 533 DIIAETPVLTOTLPEGKAHDMDMTSEVDFPFSSEAVTATETSEALRTEEVTAASGAEEET 592
 Qy 960 DMVSAVSQUTLSDPTTEATPVQVEVSGVLDTEEBERQOAILLOAVADVKESEVPATQ 1039
 Db 593 DMVSAVSQUTLSDPTTEATPVQVEVSGVLDTEEBERQOAILLOAVADVKESEVPATQ 652
 Qy 1040 TVQRTGSKALEKVEEVEDSEVLASEKEKDVMPKGPVQAGAEHLAQGETSQATPESLE 1099
 Db 653 TVQRTGSKALEKVEEVEDSEVLASEKEKDVMPKGPVQAGAEHLAQGETSQATPESLE 712
 Qy 1100 VPEVTADVHVATTCQVYIKLOQIMEQAVAPESSETLTDSETNSTPLADSDTADGTQODET 1159
 Db 713 VPEVTADVHVATTCQVYIKLOQIMEQAVAPESSETLTDSETNSTPLADSDTADGTQODET 772
 Qy 1160 IDSOSKATAAROSQVNEEBEAAATQKEEPTLPNNVPAQEHGEEPRGDLVLEPTQOEELT 1219
 Db 773 IDSOSKATAAROSQVNEEBEAAATQKEEPTLPNNVPAQEHGEEPRGDLVLEPTQOEELT 832
 Qy 1220 AAAPVLAKEVEGOEVDWMLDGEKVEEBOEVFVHSGPNSOKAADVTYDSEVGAACQE 1279
 Db 833 AAAPVLAKEVEGOEVDWMLDGEKVEEBOEVFVHSGPNSOKAADVTYDSEVGAACQE 892
 Qy 1280 KESTEVQISLSEEGMETVVEKRETKPEQVSEBGOETAPRBNHGTVGKVLTLDMSS 1339
 Db 893 KESTEVQISLSEEGMETVVEKRETKPEQVSEBGOETAPRBNHGTVGKVLTLDMSS 952
 Qy 1340 SERGKALGSLGSPSLPDDOKAGCIEVOVQSLDTTQVTAEBNEKIEFTVVISSETSESPE 1399
 Db 953 SERGKALGSLGSPSLPDDOKAGCIEVOVQSLDTTQVTAEBNEKIEFTVVISSETSESPE 1012
 Qy 1400 CVGAHLPLPAEKSSATGHTLQHAEDTVPLGPESQAESIPITVPAEESTLHPDLQGEIS 1459
 Db 1013 CVGAHLPLPAEKSSATGHTLQHAEDTVPLGPESQAESIPITVPAEESTLHPDLQGEIS 1072
 Qy 1460 ASQRESEEDRPDAGPDADGKESTAIKLVKAEPELLEESKSNKIYANVITQAVDQPA 1519
 Db 1073 ASQRESEEDRPDAGPDADGKESTAIKLVKAEPELLEESKSNKIYANVITQAVDQPA 1132
 Qy 1520 RRETAPEHTAYVSOVPAICRLDSREPNNCMTKMDAKMKHPVOPORREDLOVLTVEAM 1578
 Db 1133 RRETAPEHTAYVSOVPAICRLDSREPNNCMTKMDAKMKHPVOPORREDLOVLTVEAM 1191
 RESULT 2
 JMW0057
 C:Species: Homo sapiens (man)
 C:Date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 08-Oct-1999
 C:Accession: JMW0057
 R:Sato, N.; Kokame, K.; Shimokado, K.; Kato, H.; Miyata, T.
 J: Biochem. 123, 1119-1126, 1998
 A>Title: Changes of gene expression by lysophosphatidylcholine in vascular endothelial C
 A:Reference number: JMW0057; MUID:98269042; PMID:9604001
 A:Accession: JMW0057
 A:Molecule type: mRNA
 A:Molecule type: mRNA
 A:Residues: 1-1684 <SAT>
 A:Cross-references: DDBJ:AB003476; NID:G2081606; PIDD:BAAL1927.1; PID:di020716; PID:G208
 C:Comment: This protein regulates cell growth.
 F:433-439/Region: nuclear location signal
 F:522-527/Region: nuclear location signal
 F:591-596/Region: nuclear location signal
 F:671-676/Region: nuclear location signal

Matches 903; Conservative 179; Mismatches 378; Indels 89; Gaps 30;
 Qy 89 VQGRSEEDVREKRDVEEMANSTAVEDITTKDQGEETSEIIQIIPASENNVEWQPAESQ 148
 Db 9 VQGRSEEDVRSKDDKEMATKSAVAVHITDDQGEETSEIIQIIPASSSNNIELTQPTESQ 68
 Qy 149 ANDVGFKKVFKEVFPKFKVKKDKNEKSPVQLTLVKKDEGGAGASVAGAGHQPVSVETA 208
 Db 69 ANDVGFKKVFKEVFPKFKVKKDKNEKSPVQLTLVKKDEGGAG--AGAGHKPSTL--G 123
 Qy 209 VGESASKESEELKOSTEKOEGTLKOESSTELPLQAESDQAAEEAKDEGEKQEKPEPTS 268
 Db 124 AGPAASKESEKOSTEKEHETLKREQSHAEISPPAESQAA--ECKEKEGEKQEKPEPTS 182
 Qy 269 PESPSVNSSETSSFKKFTTHAGAWRKSTSFKKSKEEDILETAKRKEQAEKVDSEEK 328
 Db 183 AEPSTPVSSETSGTFKKFTQGWAGWRKSTSFKPKXEDVEASEKKEQEKPEKVDTEED 242
 Qy 329 EKTPEASE-----QEPADTDQARLSADYEKVELPLEDQVGLDEASEKEKAPLA 379
 Db 243 GKAEVASEELTASBOAHQPEPAESAHEPRLSAEYKVELPEBEQVSGSQGEKPEKAPLA 302
 Qy 380 TEVPDEKKEAHQ--EVVAEVAHSTVEKTEEEOGGGGEAGGVVVGTSGLPPEKLAEPQE 438
 Db 303 TEVPDEKTEVHQEEVVAEVAHSTVEBERTEQ-----KTEVEETAGSVABELVEMDA 354
 Qy 439 VPQEAPEEELMKSRMVCVSGDHTQTLTDSLPEEKTLPKABEGIVSEVMLSQERIKVQ 498
 Db 355 EPOEAPEAKEVLTKETCVSGEDPTQAGADLSPDEKVLSPKEGVVSEVEMLSQGRMKVQ 414
 Qy 499 GSPKLKLFSSGKLKLSKKKQKRGCGGDEEPGEYQIHHESPESADQEGESASAPPE 558
 Db 415 GSPKLKLFSTGKLKLSKKKQKGR--GGDEESGHHQVPADSPQSEQGEBSASSAPPE 473
 Qy 559 EPEETTCLEKPLLEAPQGEAEEGTTSDEKKRREGITPMASFKKAVTTPKKKVRPSESDX 618
 Db 474 EPEETTCLEKGLABVQDQGEAEBAITSDEKKRREGITPMASFKKAVTTPKKKVRPSESDX 533
 Qy 619 EEELEKVKSAITLSDTSVSEMQDEKTVGEBQKPEEKKRRVDTSVSWEALICVSSSKR 678
 Db 534 EDELDKVKSAITLSTESTASEMGEEMKGSVEBPKEEPRKRVDTSVSWEALICVSSSKR 593
 Qy 679 ARKASSDDEGPRRLTGGDSHRAEASDKDAGTDAVAVASQEDQAGSSSPAPASPS 738
 Db 594 ARKGSSEDEGPKAMGDDHQKADGAKETGTGDIAGQEHDPGGSSSPAPAGSPPT 653
 Qy 739 EGEGSTVESFRLVTPPKKSKSKLEEAED---SSVEQLSTELIEPRRESSWVSIKKEI 794
 Db 654 EGEGSTVESFRLVTPPKKSKSKLEESSEDSIAGSVGHESTPTEPRKESWSVSIKKEI 713
 Qy 795 PGRKKRADGKQEQATVEDSGPVEINEDDPNVPAVVPPLSEYNAVERKME--AQGNTL 851
 Db 714 PGRKKRPDQKQEQAPVEDAGPTGANEDSDVPAPVPLSEYDAVERKMEAGQAQKSAEQ 773
 Qy 852 POLGAVVSEELSKTLVHTVSVAVI DGT RAVTSVEERSPSWI SASYTEBLEHTAGAMP 911
 Db 774 POLKATEVESKELSEQVHMAAAVADGTTRATTTIIEKSPSWISASYTEBLEHTAGAMP 833
 Qy 912 PVEETEKDIIA--EETPVLTOTLPEGKAHDMDMTSEVDFPFSSEAVTATETSEALRTEEV 970
 Db 834 LTEEVLESEVLAEEPPVTEPLPENRARGDTVASEALPTEAVTAAETAGPLCAEBCGT 893
 Qy 971 EASGAETTDWVSAVSQUTLSDPTTEATPVQVEVSGVLDTEEBERQOAILLOAVADV 1030
 Db 894 EASGAETTDWVSAVSQUTLSDPTTEATPVQVEVSGVLDTEEBERQOAILLOAVADV 953
 Qy 1031 EESQVPAT-----QVQRTGSKALEKVEEVEDSEVLASEKEKDVMPKGPVQAGAEH 1083
 Db 954 EESQVPATGPGGPEVQLPQVRA-----EAEREDEQAEASGLKKEEDVVLKDAQGAKEP 1007
 Qy 1084 LAQGETSQATPESLE--VPEVTADVH---VATCOV-----IKLQOL--MEQAVAPESSET 1133
 Db 1008 FTQGVAVGQTTPESEKAPQVETESLESSELVTTQAEHTLAGVKSQEMMEQAIIPDSVET 1067

Query Match 50.6%; Score 4083.5; DB 2; Length 1684;
 Best Local Similarity 58.3%; Pred. No. 5.6e-139;

QY 1134 LTDSNTNGSTPLADSDTADGTOQDETIDSQDSKATAAARQOSQVTEBBAATAQKEBETSLP 1193
Db 1068 PTDSGTGTPVADFDPAGTQDDEIVIEIHEENEVASGTQSGGTEABEAVPAQKERPPA-P 1126
QY 1194 NNPVPAOEHEEFG-RDVLPTQOELTAAAPVVLAKTEVGEVDMLDGEKVK- - - - -E 1247
Db 1127 SSFVQETKEQSKMEDTLEHTDKESVETVSIUSKTEGTQ- - - - -EADQYADEKTKDVPFFE 1184
QY 1248 EQEVFVHSG- - - - -PNSQKAADVTYDSEVMGVAGCQEKESTEVQS- - - - -LSLEEEMETDVE 1300
Db 1185 GLEGSIDTGTIVSREKVTVALKGEGETEAECKKDDALELQSHAKSPSPVEREMVQVE 1244
QY 1301 KEKRETKPEOVSEEG- - - - -EQTAAAEHEBGTGKPVLTLDMPSSRGKALGSLGGS- - - - -PSLPDQ 1358
Db 1245 REXTEAEPHTVNEKLEHEHTAVTVSEEVSKQLLTQVNPVITDGAKEVSSLEGSPPPCLGQ 1304
QY 1359 DKAGCTEVQVQSLDVTVTQTAABAEKVI- - - - -ETVVISSTGSPRCVCGAHLPAEKSSATGG 1416
Db 1305 EBAVTKIQVQSSEASTLTAAABEEKVLGETANILETGETLPAGAHVLLEKSSKNE 1364
QY 1417 HWTLOHAEDTVLPGPSQAESIPITVPAPESTLHPDLOGEISASQSRSEEBEDKPDAGP 1476
Db 1365 DPAAHPCEDAVPTGDCQAKSTPVIVSATTKGLSSDLEGEKTTSLKWKSDDEVQVACQ 1424
QY 1477 DAGKSTAEIKVLKAPPE- - - - -ILEBSKSNKIVLNIQTAVDQFART-ETAPETHAYDSQ 1533
Db 1425 EV- - - - -KVSVAIEDL- - - - -EPENGILLETKSSKLQNIQTAVDQFVTEATATEMLTSELQ 1479
QY 1534 TVPACELDSREPNCWTK- - - - -MKDAKMKHPVPQPREDLQVLVLEA 1577
Db 1480 TQAHVTKADSQDAGQETEKEGEBPLASAOQDETPTSKEESESTAVGQA 1528

RESULT 3
T24583
hypothetical protein T06D8.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24583
R:Palmer, S.
submitted to the EMBL Data Library, April 1995
A:Reference number: Z19909
A:Accession: T24583
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1829 <WIL>
A:Cross-references: EMBL:Z491130; PIDN:CAA88964.1; GSPDB:GN000020; CESP:T06D8.1
A:Experimental source: clone T06D8
C:Genetics:
A:Gene: CBSP:T06D8.1
A:Map position: 2
A:Introns: 1391/3; 1432/3; 1470/3; 1505/1; 1520/1; 1616/1; 1644/1; 1687/3; 1742/1

Query Match 7.3%; Score 585.5; DB 2; Length 1829;
Best Local Similarity 23.2%; Pred. No. 1e-13;
Matches 375; Conservative 174; Mismatches 662; Indels 405; Gaps 65;

QY 2 GAGSSTE- - - - -QRSPEQAGSTPSELVLSHGHPAAEASG- - - - -AAGDPADADPAT- - - 49
Db 106 GSGETTVVAVVSSGSEEPASSSTSVPTLSKDDQVTEASGETTTTAAATEASSEETTTTS 165
QY 50 KLPQKNGQLSSVNGVAB- - - - -QGDVHVQENQEGQEE- - - - -VVDSDVQ 91
Db 166 AVTEGSEETTTTAVTEASSEATTTAGTASGEETTTTAVTEGSGEETTVVAVVSSG 225
QY 92 RESE- - - - -DVREKDRVEMAANSTAVEDITKDGQETSEIIIEQIPASENNVEKVV- - - 142
Db 226 EPASSSTSIPTLSKNDQVTEASGEETITAAATEASEETTTTAVTEGSGEDTTVAVVEL 285
QY 143 - - - - -OPAESQANDVGFKVFFVGFKFTVKDKNEKSDTVOLLTVKDDGEGGAASVAG- 198
Db 286 SGQAPASS- - - - -TSIPTLSKDDQV- - - - -TEASGEETTTAAAT 321

QY 199 DHOEPVETAVGESASKESELKQSTBKQGTULKQBOSSTEIPLQAESDQAAEEAEKDEGE 258
Db 322 EASEETTSVAVTEGSGEETTVVAVVSSG- - - - -EPASSSTSIPTLSKDDQVTEAS- - - - -GE 376
QY 259 EKQEKPTK-SPESPSPV- - - - -NSETTSFKFFTHGWAGWRKKTFFKSKEDDLETAEK 314
Db 377 ETTAAATEASEBETTSVAVTEGSGEDTTVAVVSSGEPASSSTSIPTLSKDDQVTEA 436
QY 315 RKQEAEKVDEEKEKETEPASBEEQPAEDT- - - - -DQARLSADYKVKVLEPLDQV 364
Db 437 SCBETTTAAATEASEBETTSVAVTEGSGEDTTVAVVSSGEPASSSTSIPTLSKDDQV 496
QY 365 GDLAASSEKCLAPLATEVDFDERKMEAHQEVVAEVHVHSTVKTETEEGGGGEAEAGVVEGT 424
Db 497 - - - - -TEASGEETTTAAATEASEE- - - - -TTTSAVTE- - - - -GSGBETTVVAVVSS 538
QY 425 GESLPPEKLAHQEVPOEAEPAEELMKSRMCVSGDHT- - - - -QLTDLSPEEKTLPRHPEG 481
Db 539 GE- - - - -EPASSSTSIPTLSKDDKVTASGEETTTAAATDASSEETTSVAVTEG 588
QY 482 IVSEVEMLSQERIKVQSGPLKPLKLFSSSGLKLGKKQKGRGGGGDEPPGEYQ-HIHT 540
Db 589 SGEETTVAVVE- - - - -SSDEEPASSSTSIPT 616
QY 541 SPESADEQKGESS- - - - -ASQPEEPEETTCLEKGPLAPQDGEAEETGTT- - - - -SDGEK 589
Db 617 L- - - - -SKDDQVTEASGEETTTAAATEASEETT- - - - -TSAVTEGSGEETTVVAVVSSGEE 668
QY 590 KRIGITPWLSPFKMVTPKVRVRPESDKBELEKVKSATLSTDTSTMVEMQDEVTGVE 649
Db 669 - - - - -PASSSTSIPT- - - - -ELSKDDKVTASGEETTTAAATDASSEETTTSAVTEGS 716
QY 650 - - - - -EOKPREPKRRVDTSVSWEALICVGSKKRARKASSDDDEGPRTLGDSH 699
Db 717 GBETTVVAVVSSDEEPASS-STSIPTEL- - - - -SKDQVTEASGEET 759
QY 700 RAEASDKKEAGTDAVPASTQEQDQAQ- - - - -SSSPBPAGSPSEGEVSTWESFKRLVTPR 756
Db 760 TAAATEASEBETTSVAVTEGSGEETTVVAVVSSGEEBPASSST- - - - -IPT 805
QY 757 KKSCKLEKAEADSVVEQLSTEIPEPSRESVSIKAFIPGRKKRADGKQEQATVDSGP 816
Db 806 ELSKD- - - - -DQVTEASGEETTTAAATEASEETTS- - - - -AVTEGSG- 843
QY 817 VEINEDDPNVPVAVPULSEYNAVEREKMAQGNTELPOLLGAVVSVSELSKTLVHTVSVAV 876
Db 844 - - - - -EDTTVAVVSS- - - - -GEQPASSSTSIPT- - - - -TELSK- 872
QY 877 IDGTRAVTSVEERSPSWISASVTEPLEHTAGEAMPPEVVEVTEKDIIAETPTVLTLPEG 936
Db 873 -DDQVTEASGEET- - - - -TAAATEASEETTTSA- - - - -VTEGS- - - - -GEETTVVAVVSSG 919
QY 937 KD- - - - -AHDDMVTSEVDTSEAVTATE-TSEALRTEEVTEASGEETDMVS 983
Db 920 EEPASSSTSIPTLSKDDQVTEASGEETTTAAATEASEBETTTTSAVTEGSGEETTTSAVT 979
QY 984 AVSOLTDSPTTTEATPVQEVSGVLDTTEEBERQTCAILQAVADKVK- - - - -EESQVPATQ 1039
Db 980 EGS- - - - -GBETTSVAVP- - - - -EGENSTTEAPAFVTGSEIEIPSSSESSSTTH 1024
QY 1040 - - - - -TVQRTGSKALEKVEEVEDSEVLASEKEKDVMPKGPVQAGAEH- - - - -LAQ 1086
Db 1025 DRSIPVITPKPSVSSITENVMS-KTSSSEAAEKIIGEHTQKDDAGKEDENMFAVTV 1083
QY 1087 GSETGQATPESLEVPVETADVHVATCQVIKLQQLMEQAVAPESSETLTDSETNGSTPLA 1146
Db 1084 ANPAGTSTTESAENVSTGEED- - - - -ENIKMAKELGKQFAAD- - - - -LA 1123
QY 1147 DSDTADGTOQDETIDSQDSKATAAARQOSQVTEBBAATAQKEBETSLPNNVPAQEEHGP 1206
Db 1124 KLAADKGVNLTADAKDSGETAHVEDEQVSTSESSIGSEETTTTV- - - - -NKETTEHHEAS 1181

QY 1207 GRDVLPEPTQOELTAAAVPLATKEVQEGEVDVMDGEKYEBOEVFVHSGSPNSOKADYV 1266
 Db 1162 G-----EEDDAPFPVTGAPLDPSTTEASVSTSA--ITTDETSVAADBESTSTSGEYQ 1231
 QY 1267 YDSEVWGVA--CCQEKESTEVOSLSLEEGEMETDVEKEKRETPKEOVSEEGEOETAAPEH 1324
 Db 1232 SSSAIIIDSTVASEBPTSEATSIVLESSE--EYTTTDEMLVSTVAQLEGGSGITAAESK 1290
 QY 1335 EGTYGKPVLTLMPSSEKALGSLGGSPS--LPDODKAGCIEVVOYOSLD---TTVTOT 1378
 Db 1291 D----EDSVTTEATSGSTTVSSSDSGESTVAPNDSESTSTESSQSTDESGVTAAES 1346
 QY 1379 AAVEVLE--IVVISSETESEPCVAHLLPAAKSSATGGHMTLQAAEDTV---PLGPES 1433
 Db 1347 KHEESSTTAPAPVNTSKITGSE--DEEDSPTEHFTIGIDETMFKSLIVPT 1397
 QY 1434 QAESIPITTPAPESTLHPDLQGEISASQERSEEDKPDAGDADGKESTAIEKY 1489
 Db 1398 HREDLPNVGVFPSPSEPKKNDE-----EEBEEBDGTKSDYEDNVSKTI 1444

RESULT 4

T13564
 microtubule-associated protein homolog - fruit fly (Drosophila melanogaster)
 N/Alternae names: hypothetical protein EG:4984.1

C/Species: Drosophila melanogaster

C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C/Accession: T13564

R/Spinos, L.; Papadogiannakis, G.; Siden-Kiamos, I.; Louis, C.

submitted to the EMBL Data Library, April 1999

A/Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A/Reference number: Z17689

A/Accession: T13564

A/Structure: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-5327 <SPA>

A/Cross-references: EMBL:AL031128; PIDN:CAA20006.1

C/Genetics:

A/Cross-references: FlyBase:FBgn0025392

A/Introns: 24/2; 52/3; 104/3; 179/1; 232/1; 1669/3; 2566/1; 4798/3; 5272/1

A/Note: EG:4984.1

C/Superfamily: Drosophila 576k microtubule-associated protein homolog

Query Match 7.1%; Score 574; DB 2; Length 5327;
 Best Local Similarity 21.6%; Pred. No. 9.2e-13;
 Matches 415; Conservative 285; Mismatches 734; Indels 484; Gaps 82;

QY 6 STHQRPBPQAGSDTSELVLSGHGPAEASGAADPAD-----ADPATKLPKNGQL 58
 Db 2744 SYVESKDKDAEKSESRESPIVAGPEVPRESKSLDPSKDTSRGVSVEYTADEKSBQ 2803
 QY 59 SSVNGVAE--QGDVHVOENQOGEVEVDEVDGQRESEDEPREKDRVEVA--ANSTA 112
 Db 2804 SRRESVAESVKADTKDKGSGEASRPSVDL---KDDDEKESRQSTTSGHKAMSTM 2860
 QY 113 VEDITDGOEETSEITIEQIPASNNVEEMVQ-----PAESQANDV--- 152
 Db 2861 GDESPMDKADSKKE-----PSRESVAESIKHENTYKDEESPLGSRDVSVAESIKSDITKG 2915
 QY 153 -----GFKYVFKVGVFKYTKDKDKNEKSDTVOLLTVYKQEGEASVAGDQEP--- 203
 Db 2916 EKSPPLSKESVPRESVVGSIKDEKASRRRESVAESVYPESKDATAPPEKESHPRESVL 2975
 QY 204 -----SVETAVGEGASKESEIKOSTE-----KQEGTLKQEOSTEIPL 241
 Db 2976 GSLKDEGDKTTSRRVAVADIKDEKSLIVSQEASRPESSEASLKDAAPSOEHSRRESVA 3035
 QY 242 QAESD-----QAEEEAQDEGEKQEKPTKSPESPSPVNSSETTSFKKFP 288
 Db 3036 ESVKDKSPVASKASRPASVAENAKDADSEKORPESIPQSKAGSIKDEKSP----- 3089
 QY 289 THGMAGMRKKTSPKSKED-----LETAERKEQ----- 318

Db 3090 -----LASKDEAKESKESRESRESVAEQPLVSKESRNPASVAESVKDEAKESKEEPLM 3143
 QY 319 -----EAKYDEEEK-----EKTE-PASEOEPAEPTDQRLADYKY 356
 Db 3144 SKASRPASVAGSVKDEAKESKESRRRESVAEKSPLPKESRPASVAESVKDEADSK 3203
 QY 357 ELPLEDQGDLEASSEKCAE--LATEVFDE-----KMAHOEVAEVA----- 398
 Db 3204 ESRRSGAEKPLPASKESRPASVAESIYKDEAKESKESRRRESVAEKSPLPKESRPTS 3263
 QY 399 -----VSTVEKTEBEOGGGAEAGVVEGTGESLPPEKLAEPQEPQEAPEELMKSR 453
 Db 3264 VAKSVKDEAKESKESRRSDVAEKSPPL--ASKASRPASVAE--SVQDEAKSK--ESR 3317
 QY 454 EMCVSGGHTQTLTDSPEEKLTPKH--PEGIVSEVEMLSQORIK--VQSGPLKULFSS-- 508
 Db 3318 RESVA-----EKSPLAYKEARPPASVAESIDEAKESKESRRRESVAEKSPASKASRP 3372
 QY 509 -SGLKLLSGKKQKRGGGGDE--EPGEYGHITESPESADQEGESASSPPEPEPTTC 565
 Db 3373 TVASVVKDEAKESKESRRSDVAEKSPASKASRPASVAESVKDEAKESKESRRRESV 3432
 QY 566 LEKGPL---EAPQDEAEEGTTSDGEEKREGITTPASFKMVTTPKRVRR----- 613
 Db 3433 AEKSPASKASRPASVAESVKDEAKESKESRRRESVAEKSPASKASRPASVAESVKD 3492
 QY 614 -SESDKEE--ELEVKASATTSSTDVTSSEMODEVYTVGEQKPEEPKRRVDTYSWEAL 669
 Db 3493 EAEKESKESRESRESVAEKSPPLSKASRPASVAESVKDEAKSK--EESRRE--SVAEKSS 3548
 QY 670 ICVSGSKRARKASSDDEGGRTLGDSHRAEBSKQKEXGTDAV--PASTQO--OPDAQ 726
 Db 3549 IASKASRPASVAESVKDEA--EKSKESRRRESVAEKSPASKASRPASVAESVKDEAE 3606
 QY 727 GS-----SSPEAGSPSEGGVSTWSEFRKLTVPKRSKSLKEAEDSSV--- 772
 Db 3607 KSKEVSRRESVAEKSPPLSKASRPASVAE--ESVVD--EADSKESRRRESGAEKPLASM 3663
 QY 773 -----EQLSTIEPSREBS--WVSIKFTIPGRKKR---ADKQEOA--TVEDS 814
 Db 3664 EASRPASVAESVKDETEKSKESRRRESVTEKSPPLSKASRPASVAESVKDEAKESKEES 3723
 QY 815 GPVELNEDP-----NVPAYVPLSEYNAVERREKQEAQNTLPLLAIVYSEELSTL 868
 Db 3724 RRESVAEKSPASKASSRPASY-----ASIKDEAGTQO-----EERRESM 3765
 QY 869 VHTVAVAVIDGTR--AVTSVERSPSMISASVTEPLEHTAGEMPPVEVTEKDIITA--- 923
 Db 3766 PESGKAESIYKQDQSLAKETSRPDSVVESVYKDETEKEGSAIDKQVASRPESVAEAK 3825
 QY 924 -ETTVLQTLPEGKDAHDNMTSEVDTSEAVTATETSEALRTIEVTASGAETTMV 982
 Db 3826 DEKSPL--HSRPE-----SVADKSPDA--SKASRSLSVAE--YASSPIIEGCP--- 3867
 QY 983 SAVSQTLTSDPTTEEA-----TPVOVESGVLDTEEEHQOTALLQAVADKYKESQ 1034
 Db 3866 RSIADLSPLNLTGAKGLPTLSSPIDVABGDFLEVAAESSPPRAVLSPKPEFSQPTG 3927
 QY 1035 VPATQVORTGSKALEKEVEVEDESE---VLASEKEDVMPKGVQVQAGAEHLAQSET 1090
 Db 3928 HTASTPVE--ASPVLEIEIVDQHTSGVGATGAFAFDL---DLTETKSTTVYKQET 3983
 QY 1091 G-----QATPSLE-----VPEYADVAVATCQVILQQLMEQAVAPESSETLT 1135
 Db 3984 TLFTLTLSKVESKAVLSSSVQVVEKQVTSYKQAEITVTLTSLQTLT-----KSEQLT 4038
 QY 1136 -----DSETNSTPLADS--DTADGTQODERTIDSDQSKATVAVQSQVTEBEATPAQK 1187
 Db 4039 EIKSVLDTNISVNTLFTAVETIEKKVQDVT--EKVIEKATIEHVSHTTGTGESSSTESQ 4097
 QY 1188 EPSTL-----PNNVPA-----QEEHGEPEGRDV--- 1210
 Db 4098 EKSSLDLTGFSRLRETHITTVGSPPEFTVYICGRDPLVLAHDIKEDEBHRFPSPVDKXA 4157

```
QY 1211 -----LEPTQOELTAAAPVLAKTEVQGEVDWLDG-----EKVKEQEVFVH 1254
Db 4158 IIPQPMRPLSPREEEV-AKIVADVAKVLKSDKDITDIIPDFBERQLKEELKSTADTEEE 4216
QY 1255 SGPNs--QKAADVTYDSEVMGVAGCQKESTEVQS--LSLEEGEMETDVEKEK-----R 1304
Db 4217 SDRSTRKSLKSVKVEIE-----SEKSSPDQKSGPISIEEKDKIQSEKAQLRQIGILA 4271
QY 1305 ETPEQVSEGE-----QTAPEHEGTYGKPVLLTDMPSRGRKALGSLGGSPSLPDQ 1358
Db 4272 SSRPESVASQSPSPSQAASHEHK-----EVELSESHKAESK-----SRPES 4316
QY 1359 DKAGCIEVQOSLDTTVTQTA-----EAVEKVIETVWISSETGESPECVGAHL 1405
Db 4317 -----VASQVSEKDMKTSRPSASTSQSTKEGDEETESLLHSLTTTETVETQMBEKS 4371
QY 1406 LPAEKSATGGHWTLOHAEDTVPGLPSQAESIPPIIIVTAPAPETLHPDLQGEISASQRE- 1464
Db 4372 FESVTSVTKS--TVLSSQSTVQLREESTESL-----SSSLKVE-----DSRRRES 4416
QY 1465 -RSEEDKPDAGPDAGKESTAEIKVLKASPEILELESKNKIVLNVIQAVDPQAFARTE- 1522
Db 4417 LSSLAEKGGIATNTSLKEDTS-ASASQLEELLVQSECSSESISVBIQTSIAQKSNKEI 4475
QY 1523 -TAPETHAYDSQTVQACRLDSREPNCRTKMDAKMKHPVQPPREDLQVLTVLEAWA 1579
Db 4476 KDARET-----KVTSQFTTTTSATKDDSLKTVABFLATEKIVSAKEAFS 4521

RESULT 5
QRNSP1
microtubule-associated protein MAP1b - mouse
N:Alternate names: microtubule-associated protein MAP1(X); microtubule-associated protein
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text_change 01-Sep-2000
C:Accession: S07549; S44387; A33645
R:Noble, M.; Lewis, S.A.; Cowan, N.J.
J. Cell Biol. 109, 3367-3376, 1989
A:Title: The microtubule binding domain of microtubule-associated protein MAP1b contains
A:Reference number: A33645; MUID:90094539; PMID:2480963
A:Accession: S07549
A:Molecule type: mRNA
A:Residues: 1-2464 <NO>
A:Cross-references: EMBL:X51396; NID:G52999; PIDN:CAA35761.1; PID:G53000
R:Sanchez, C.; Padilla, R.; Paciucci, R.; Zabala, J.C.; Avila, J.
Arch. Biochem. Biophys. 310, 428-432, 1994
A:Title: Binding of heat-shock protein 70 (hsp70) to tubulin.
A:Reference number: S44387; MUID:94234720; PMID:8179328
A:Accession: S44387
A:Status: preliminary
A:Molecule type: protein
A:Residues: 653-663, 'IC' <SAN>
C:Superfamily: microtubule-associated protein MAP1B
C:Keywords: microtubule binding; phosphoprotein; tandem repeat
F:589-786/Domain: microtubule binding #status experimental <MTB>
F:589-592,639-642,649-652,655-658,660-663,668-671,674-677,679-682,683-686,687-690,691-69
R-K-E/D-X)
F:1861-2064/Region: 17-residue repeats
F:91,116,351,888,1124,1153,1168,1208,1662,1877,1918,2003,2030,2054,2083/Binding site: ph
F:147,969,1336,1562,1563,1702,1708,1990,2057,2063,2419/Binding site: phosphate (Thr) (cd
F:1953/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match
Best Local Similarity 21.1%; Pred. No. 1-5e-10;
Matches 307; Conservative 266; Mismatches 648; Indels 532; Gaps 86;

QY 45 ADPATKL-----PQKQQLSSVNGVAFQGDVH-----VQEEQEQEEVWDVGYQRE 93
Db 478 ANPAEKIIRVFPGNSTQVNILEGLEKHLKHLDFLKQPLATKQKLTGOVPTPPVQVKLQK 537
QY 94 SEDVREKDRVEEMAANSTAVEDITKQGBETSEI-----IEQIPASENVEEMVQPAESQ 148
```

```
Db 538 RADRSRES---LKPATKPVASKSVRKESKEETPEVTKTSQVEKTPKVESKE----- 584
QY 149 ANDVGFKVKFVGFKFTVKDKNEKSDTVQLLVKDKDEGEAGBASVGAGDHQEPSEVETA 208
Db 585 -----KVLVKDKPKVKTESKPSVT-----EKEVSKSEQSP-VKAE 619
QY 209 VGESASKSELKOSTEKQSGTLKQEOSSTETIPLQABSDQAAEBEAKDEGEKQEKPTKS 268
Db 620 VAEQKATESKPKVTKDK---VVKKE-----IKTKLEEKKEEKPKKEVVKKEDKTPLLK 669
QY 269 PRSP--SSPVNSETTSFKKFFTHGMAGWRKKTsfKKSKEDDLLETKAEKQEAEEKVDEEE 327
Db 670 DEKPRKEEVKKEIKKEIKK-----EERKELKKEVKKETPLKDAKKEVKKEKK---EVK 720
QY 328 KERTEPASBEQPAEDTDOARLSADYKVELPLEDOVGDLSESEKCAPLATEVEFDE-- 385
Db 721 KEEKEPKKEIKKISKDKKSTPQSDTKPS-ALKPKVAKKEESTKKE--PLAAGLKDKG 777
QY 386 -----KWEAHQEVVAEVHVSTVEKTEEBEOGGGEAGGVVVEGTGESLPPEKLAERQEV 440
Db 778 KVKVIKKEGKTTAAATAVGTAAATTAHVAAAGIAASGPVKELEAER---SLMSSPEDLT 834
QY 441 QBAEPAAEELMKREMCVSGGDHTQLTDLSPKEK---TLPKHPEGIVSEVEMLSQSERIKV 497
Db 835 KDFEE---LKAEEDVAKDIKPQLELIEDBEKLEKETQGEAYVIOKETEVSKGSAB--- 887
QY 498 QGSPLKLFSSGLKSLGKQKGRGGDBEPGEYQHHTHTEPSADSEKQESSASSP 557
Db 888 --SP-----DEGITTTEGE-----GECEQTPPELEPVKEQGVN--DIEKFEDEGAG 930
QY 558 EPEETTCLEKPLEAPQDGEABEGTTSQGEKKREGITPWAFKQMTPKKRVRRSESD 617
Db 931 EESSET---GDYBEKARTEAEPEEDGEDNASG-----SASK-----HSTPDED 972
QY 618 KEEELB---KVKSATLSSTSTVSEMDEVKTVGCEOKPEEPKERVDTSVSWEALICV 672
Db 973 ESAKAADVHLKESKESVVGDDRAEDMDVLEKEAQSEEGEEED----- 1021
QY 673 GSKKRAKASSSDEG-QPRTLGGDSHRAEBASKDKAGTDAVPASTQEQOAOQSS-- 729
Db 1022 -----KAEDAREEGYEPDKTEAEDYVMADVADAAEAGV-----TBEQYGLTSAK 1067
QY 730 -----SPEPAGS-----PSEGEVSWTESPK-----RLVTPRKKSQKLEBAED 769
Db 1068 QFGIOSPSREPASIHDETLPGGSESEATASDEENREDQPEEFATSGYTQSTIETSESP 1127
QY 770 SSVQQLST-----EIPSRSEESWVSIKKF-----IPGRR 798
Db 1128 TPDENSTPRDVMDETNNETESQEFVNITKYESSLYSQEYKPAVASFNGLSEGSK 1187
QY 799 KKRADGQKQATVEDSGPVVEINEDDPNPVAVVPLSGEYNAVEREKMAEQNTPLQLLGAV 858
Db 1188 TDATDGKYNASASTISPPSSMEEDKFKSA--LRDAYCSEKELKASAELEIKD----- 1240
QY 859 YVSEELSKTLVHTVSVAVJDGTRAVTSVEERSPWSIASVTEPLEHT-AGE-----AMPP 912
Db 1241 -VSDE-----RLSPA-----KGPS-LSPSPSPPIEKTPLGERSVNFSLAP 1278
QY 913 VE-EVTEKDIABEETPVLTOTLPEGKDAHDDMVTSEVDTSEAVTAT-----E 959
Db 1279 NIKVSAEGARSVPGVTVQAVVEBHCASPEKTLIEWVSPSQSVTGSAGHTPYQSPOTDE 1338
QY 960 TSEARLTE-----EVTEASGABETT----- 979
Db 1339 KSHLPLTEVSENAQVAVPSFEFESEAKDENERASLSPMDEPVPDSESPVEKVLSPLRSPPL 1398
QY 980 -----DMVSAVSQ-----TDSP-----DTTEATPVQEVES-GVLDTBEEROT 1018
Db 1399 LGSSEPYEDFLSADSKVLRRESSEPFEGKNGKQGFPPDRESVPSDLTSTGLYQDKQEKST 1458
QY 1019 QAILQAVADKVEE--SOVPATQTVORTGSKALEKVEEVEEDSEVLASEKEKDMKPGPVQ 1077
Db 1459 GFI-----PIKEDFGPEKKTSDVETMSSQSALADLERKLGDV--SPTQIDVQSGSPK 1510
```


QY 1098 LEVPEVTADVDHATCQVVKIQLQMEQAVPESSET--LTDS-----ETNGSTPLADSDT 1150
 Db 1244 -----AGPSESETQKVAADAAARKQKQETDEKQKLEAEIT 1275
 QY 1151 ADGTOODETIDSODSKATAAVROSOVTEEAATAQKEPSTLPNNVPAQEEHGEPEGRDV 1210
 Db 1276 AKSADEK-----SKLEAEKLLKAAAEVAAKQKQEKDQLKLDT---EAAKKAANEK 1326
 QY 1211 LEPTQOELTAAAVPVL-----KTEVQOGEVDMLDGEKVKQEQEVFVHSGPNSQK 1261
 Db 1327 LELEKQSHIKAAAEVDVAVKKQKLEEKQORLESEAAATKKAADAEKLLKEQ-----KKX 1378
 QY 1262 AADVYDSEVMGVAGCOKESTEVQSLSLEGEMETDVEKEKRETKPEQVSEGEQHTAA 1321
 Db 1379 AAETA-----LIEIKQEQEKLAQE--QSRLEDEAKKSAEKQKLESETSKQTEE-----A 1426
 QY 1322 PEHEGTYGKPVLTIDMPSSERKALGSLGGSPSLPDQDKAGCTEVQVQSLLDTTVO--- 1377
 Db 1427 PKE-----SVDEPKKKVLLKKTEKSDSSISQSKS 1457
 QY 1378 ----TAEAVE-----KVIVTWISETGESPECVGAHLLPAEKSSATGGHWTLQHAEDTVP 1428
 Db 1458 AKSTVDAAEETLESDFNLVEKTKVQKVEQSPD-----ESTSATIKRDPQAQTEEISK 1508
 QY 1429 LQPESQAESEPIIIVTAPESTLHPDLQGEISASQRESEEDKPDG-GPDA----- 1478
 Db 1509 QDDGDEKTKTTTGDGPKPKPEDS-----EATPKKRVVKKTKQKSDSVASDASLADVSKLS 1561
 QY 1479 DGKESATAIEKVLK-----ASPEILELES-KSNKIVLVNIQTAVDQFA 1519
 Db 1562 DDVEERPKKVLKKKTEKSDSVISSETSVDTIPKESVEIPTKRAEQMILH-----NRFS 1615
 QY 1520 RTETAPET---HAYDSQTQ 1535
 Db 1616 -TDSAVESEPKNAHKDTE 1633

RESULT 7
 S37431
 N:Ankyrin 2, neuronal long splice form - human
 N:Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid
 N:Contains: ankyrin 2, short form
 C:Species: Homo sapiens (man)
 C:Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 13-Aug-1999
 C:Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569
 R:Chan, W.
 submitted to the EMBL Data Library, September 1993
 A:Reference number: S37431
 A:Accession: S37431
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-3924 <CHA>
 A:Cross-references: EMBL:226634; NID:9406287; PIDN:CAA81387.1; PID:9406288
 R:Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
 J. Cell Biol. 114, 241-253, 1991
 A:Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a
 A:Reference number: A39643; MUID:91302466; PMID:1830053
 A:Accession: A39643
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2077 <OTI>
 A:Cross-references: GB:X56957
 A:Accession: B39643
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1443, 3585-3924 <OTT>
 A:Cross-references: EMBL:X56958
 R:Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,
 Genomics 10, 858-866, 1991
 A:Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
 A:Reference number: A40334; MUID:92009921; PMID:1833308
 A:Accession: A40334

A:Molecule type: DNA
 A:Residues: 463-474, 'PB', 477-495 <TSE>
 A:Cross-references: GB:M37123; NID:9178647; PIDN:AAA62828.1; PID:9178648
 R:Chan, W.; Kordeli, E.; Bennett, V.
 J. Cell Biol. 123, 1463-1473, 1993
 A:Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and s
 A:Reference number: A49462; MUID:94075409; PMID:8253844
 A:Accession: A49462
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-3924 <RES>
 A:Cross-references: EMBL:226634; NID:9406287; PIDN:CAA81387.1; PID:9406288
 C:Genetics:
 A:Gene: GDB:ANK2
 A:Cross-references: GDB:127607; OMIM:106410
 A:Map position: 4q25-4q27
 C:Superfamily: ankyrin; ankyrin repeat homology
 C:Keywords: alternative splicing
 F:2-3924/Product: ankyrin 2, long form #status predicted <MAT>
 F:129-161/Domains: ankyrin repeat homology <AN01>
 F:162-190/Domains: ankyrin repeat homology <AN03>
 F:191-223/Domains: ankyrin repeat homology <AN04>
 F:223-264/Domains: ankyrin repeat homology <AN05>
 F:265-297/Domains: ankyrin repeat homology <AN06>
 F:298-330/Domains: ankyrin repeat homology <AN07>
 F:331-363/Domains: ankyrin repeat homology <AN08>
 F:364-396/Domains: ankyrin repeat homology <AN09>
 F:397-429/Domains: ankyrin repeat homology <AN10>
 F:430-462/Domains: ankyrin repeat homology <AN11>
 F:463-495/Domains: ankyrin repeat homology <AN12>
 F:496-528/Domains: ankyrin repeat homology <AN13>
 F:529-561/Domains: ankyrin repeat homology <AN14>
 F:562-594/Domains: ankyrin repeat homology <AN15>
 F:595-627/Domains: ankyrin repeat homology <AN16>
 F:628-660/Domains: ankyrin repeat homology <AN17>
 F:661-693/Domains: ankyrin repeat homology <AN18>
 F:694-726/Domains: ankyrin repeat homology <AN19>
 F:727-759/Domains: ankyrin repeat homology <AN20>
 F:760-792/Domains: ankyrin repeat homology <AN21>
 F:793-825/Domains: ankyrin repeat homology <AN22>
 F:826-858/Domains: ankyrin repeat homology <AN23>

Query Match 6.1%; Score 493; DB 2; Length 3924;
 Best Local Similarity 20.4%; Pred. No. 5, 1e-10;
 Matches 422; Conservative 270; Mismatches 680; Indels 698; Gaps 99;

QY 68 GDVHVOENQEQEBEVDVQRESEVDREKRVVEEMAANSTAVEDITKQGEETSE- 126
 Db 1497 GSIKVKELVKAEEBPGPEFIVERVKEDL---EKVNEILRSCTCTRD-ESSVQSSRSR 1552
 QY 127 -IIEQ--IPASENNVEEMVQPAESQANDVGFKKFKVGFVKFKDKK----NEKSDTV- 178
 Db 1553 GLVEEWEVTVSDEEIEEAARQKAPLEITEYPCVEVRIDKEIKGKVEKDSGLVNYLTDDL 1612
 QY 179 -----QLLVTKKDEGGAESVAGDHQPSVETAVGESASKES-----ELKQSTK 225
 Db 1613 TCVPPLKEQLTVQDKAGKCEA-----LAVGRSSEKGGKIDIPDETOSTOK 1659
 QY 226 Q-----EGTLKQESSTEIPLOAESDQAAEEAAKDEGEKQEKETKSP----- 270
 Db 1660 QHKPSLIGIKKPVRRKLKEKQKQEEGLQASAKA--ELKKGSSSESLGSDPLGLAPEPLT 1717
 QY 271 -SPSPVNSSET--TSSFKKPFTHGWAGWRKKTSTFKSKEDD-----LETAERKQEQEAK 322
 Db 1718 VKATSPLEETPIGSIKD-----KVKALQKRVEDEQKGRSKLPTRVKGKEDVPKK 1767
 QY 323 VDEEEKETEPASEEQ-----EPASDITQARLSADYKVE----- 357
 Db 1768 TTHRPHPAASPLKSERHAPGSPFKTEHSTLSSSAKTERHPVPSPKTEKHSFVSPS 1827
 QY 358 LPLEDQVGDLEASSEKCAPLA-----TEVFKMEAHQEVV----- 394

```

Db 1828 AKTERHSPASSSSKTEKISPVSPSTTERHSPVSSKTRHPVPSGKTDKRPVSPSG 1887
Qy 395 -AEVH-----VSTVEKTEEO--GGGGAEGGVVEG--ESLPRKXAP 436
Db 1888 RREKHPVPSRGREKRLPVSPSGRTDKHQPVSTAGTEKHLPVSPGTEKQRPVSPSTK 1947
Qy 437 QEVPOEAPAEELMKSRKNCVSGD-----HTQULTDSPEEK-----TLPRHP-- 479
Db 1948 TERIEETMSVREIMKAFQ---SGODPSKHKTGLFEHKSAROKOPOEKGHYVEKEKGPIL 2004
Qy 480 ---EGVSVEMLSQERIKVQSGPLK---LFSSGLKKLSGKQKGRGGGDE--- 530
Db 2005 TOREAKTENQTKKQGRLPVSTASKRGVRSIGVK-----EDAGGKXVLS 2056
Qy 531 -----PGEYOHITESPES--ADEOKGSSASPEE----- 559
Db 2057 HKIPREVQSVPEEHSREBEVPEKMAADQGDMDLQISDRKSTDPSEVIKQLELDNDK 2116
Qy 560 -----PEETCLEKPLEAPQ-----DGEAE--- 580
Db 2117 YQOFLSEET---EKAQLHDQVLTSPFNTTFLDYMKDEFLPALSLQSGALDGSSESUK 2173
Qy 581 -----EGT---TSOGEKKEG--TPMASFKKM--VTPRK--RYMRPSESUK 618
Db 2174 NEGVAQSPGSLMEGTPQISSSESYHGBLAETPETSPELSFPKSKSEBQIGETESTK 2233
Qy 619 EEBLEKYS-----ATLSSTDSVSEMDVKTVGEQKPEEKREY 660
Db 2234 TETTTTIREKHEPTTKDITGSEERGATVTEDESETSESFOKEALTG--SPKOTSFKQD 2292
Qy 661 DTFSVWEALI CVGSSKKRKARKAS-----SSDDEGCPRTLGGDSHRAEASGOKAGTD 713
Db 2293 D-----CTGSCVVALAKETPTGLTEBAACDE--GORFSSSAKHTQDSEAO--- 2338
Qy 714 AVAPSTQEDQAGSSSPBA-----GSPSEBGVTSWSEFKRLVTPRKSS----- 759
Db 2339 ---STATSDETALPLPEASVKTGTGTESKPQGVIRSPQGLFELAPSRDSEVLSAVAD 2394
Qy 760 ---KSKLEEK--AEDSVQOLSTEIERS--REBSWVSIKKIFGRKKRA----- 802
Db 2395 SLAVSHKDSLSEAPVLEDNSHHTPDSLEBSPLKESPCRDLSLESSVPEPKMGKGTFSHP 2454
Qy 803 -----DGKOBQATVEDSGPVEINEDPN----- 825
Db 2455 PLPAVAKTELTTEVASVSRLLRDPDGAEDDSLQTSIMESSGKSPSPDPSSEBVS 2514
Qy 826 -----VPVV-----PLSEYNVBER--- 841
Db 2515 YEVTPTTDVSTPRPAVVIHECAEEDSENGEKKRFTPEEEMFKMTKIKWDELQEOAKQ 2574
Qy 842 -----KMAQGNTELPOLLGAVVSEELSKTLVHT-----VSVAVIDGTAVTS 885
Db 2575 KRDYKPEKQOESSSSSDPADCSVDVDE---KHTSGDESGVPIVLTSESKEVSS 2629
Qy 886 VEERSPSW-----ISASVTEPLEHTAGAMPVEEYTEKDIATETPVL 929
Db 2630 SSSSEBELAQKKGADSGLLPEPVIRVQPPSPSSMDNSNP--EEVQFQPVVSKQ---- 2684
Qy 930 TQTLPRGKAHDMVTSEVDTSEAVTATETSEALRTE-----EVTAS 973
Db 2685 -YTFKNMEDTQOEBPGKSEEEKDESHLA--EDRAVSTEADRSYDKINRDPQKICDGH 2742
Qy 974 GAEETIDWASA--VSQUTDSP-----DTTEATPVQEVESG 1007
Db 2743 GCAMPSSSARPVSSGLOPFGDDVDQPVITYKESIALQGTHEKOTEGSELVLSAESP 2802
Qy 1008 VLDTEEBERQQAIIQ--AVADKYKEESQVPATQVQRTG--SKALEKVEVEED----- 1058
Db 2803 QADCPSESFSSSSLPFCIVSEKGELEDIEDISATSSIOKTEVTXTDEFFENLPDCCSODS 2862
Qy 1059 -----SEVLASEKKEKVMKGPQOEGAEHLAAGSEFGQATP-----E 1096

```

```

Db 2863 SITQTDRFMDVPVSLDAENDEIYDQITSPYENVPQSPFSSSESKTQTDANHTTFSH 2922
Qy 1097 SLEVPBV--TADVDAHATC-----QYIKLQQLMEQAVAPSSSETLIDSE 1138
Db 2923 SSEVSVTITSPVEDVVAVSSSSGTVLSKESNFEQODIKMESOLESTL---WEMQSDSV 2978
Qy 1139 TNGSTP--LADSDTADGQODTIDSQ---DSKATAAVR-----OSQYTEER---AATA 1184
Db 2979 SSSFEPTWSATTYVVEQOISKVITIKTDVSDSWEIHEDEABEARKVEEOKIFGLMV 3038
Qy 1185 QKEEPTSLPNNVPAQOEHEEERGRDVLPTQOEL--TAAAVVLAKT---EVQGEVDM 1239
Db 3039 DRQSGTTPPTTPAR-----TPTEGTPTSQNFPLQEGKLFMTSGAID-- 3085
Qy 1240 LDGEVKEEQQVFNHSGNSQKADVTYDSVM--GVACQCKE--STEVQSLSEEGMET 1297
Db 3086 MTKRSYADESFHFQIG---QESREETLSEBDKEGATADPLPLETSAESIALSESKEV 3142
Qy 1298 DYKEREKRETKPQVSEEBEQ-----ETAAPEHGTGKPLVLTDMPSSE 1341
Db 3143 D---DEADLLPDSVEEVEEIPASDAQNSQMGISASTETTKAV---SVGTQDLPTVQ 3196
Qy 1342 RGKALGSLG--SPSLPDQKAGCIEVQVQSLDTTVTQTAEAVKVLTEVVISERG--ESP 1398
Db 3197 TGD-IPPLSGVKQISCPDSEBA---VQVQDLFSLTRS-----YSDRGDDSP 3241
Qy 1399 ECVGAILLPAEKSSATGGMWTLQHAEDTVPLGP-----ESQAEIPI---ITPAPPE 1447
Db 3242 DS---SPEEQKSVI-----EIPAPMENVPFTSKSK--IPVRTPMTSTPAP 3284
Qy 1448 STLHPDLQGEISASGRSESEEDKPDADGKSTAIKVL--KAPPEILLESKSKI 1506
Db 3285 SA---EYESVSEDLSSVDEENKAD---EAKPKSKLPVKPQLOVEQQLSDLTISVOKT 3338
Qy 1507 VL---NVLOTAVQFARTETAPETHAVDQTOVPACRLDREBPNCKTKMKADAKHAPV 1562
Db 3339 VAPQGDMASTAPDRKSES--DASSLDSKTKCV-----KRSYTERETESREK-- 3387
Qy 1563 PQRREDQVLTVLEMAQPRKCLPRLQKA 1592
Db 3388 ---AELE-LSEEGATRPKILTSRLPVKS 3413

RESULT 8
T29757
protein UNC-89 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Dec-1999
C:Accession: T29757
R:Du, Z.; Le, T.T.; Wilson, R.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid C09D1.
A:Reference number: Z20679
A:Accession: T29757
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6642 <DUZ>
A:Cross-references: EMBL:AF003131; PIDN:AB54132.1; GSPDB:GN00019; CESP:unc-89
A:Experimental source: strain Bristol N2; clone C09D1
C:Genetics:
A:Gene: CESP:unc-89
A:Map position: 1
A:Intons: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1; 6
3/3; 5917/1; 6027/1; 6061/3; 6153/2; 6515/1; 6552/3; 6609/1

Query Match 6.1%; Score 489; DB 2; Length 6642;
Best Local Similarity 21.8%; Pred. No. 1.3e-09;
Matches 348; Conservative 234; Mismatches 670; Indels 346; Gaps 68;

Qy 7 TEORSPEQAGSDTPESELVLSGHGPAEASGAADPADATKLPKXNGQLSSVNGVAE 66
Db 1132 TYQKKPFAEFLRVSLSLVEKGEAVFSAHAFGLP-----LPTYEM---SYNG--- 1177

```



```
QY 67 QGDVHVQENQGEQEEVVDVQRESEEDVREKD-----RVEEMAANSTAVED 115
Db 1178 -----RKRDQGEQGARVTRDSTVDGASILTIDTATYSEVNNHLTISVVAENTLGAEE 1230
QY 116 -----ITKDGQEBTSEITIEQIPASENNVEEMVQPAESQANDVGFKVFKVGFKF--TVKK 169
Db 1231 TGAQLTIEPKKE-SVVVEKQDLSSSEVQKIEAQOVKEASPEATTTIMETSLSTKTTM 1289
QY 170 DKNEKSDTVOLLTVKKDGGGABASV-----GAGDHQBPVSIVETAVGESAKESBLKOSTEK 225
Db 1290 STTEVTSTVGVTETKESSESATTVIGGSGGVTEGSIIVSKIEVVSXTD---SQTDV 1346
QY 226 QEGTLKQESSTPIQAESDQAEBAEKDEGEKQEKPTKPSPPSSVNSVNETSSPK 285
Db 1347 REGTPKRVFAEEELPKVIDSDRKKKSPSPDKKESPEKTEEKPASPT----- 1397
QY 286 KPFTHWAGWRKKT--SFKKSE--DDELTAEKKEQAEKV---DEBEKEKTPASEEQE 339
Db 1398 -----KKTGEVKSPEKSPASPTKESPAAEVVKSPPTKKEKSPSPPTKKEKS 1446
QY 340 PAEDTOARLSADYKVELPLEQVGDLEASBEKCAPLATEVDFDEKMEAHQBVVAEVHV 399
Db 1447 PSSPT---KKTGDEVKESPPKSPPT--KKEKSPKPEKDVKSPVKKEKSPDATNIVEVSE 1501
QY 400 STVEKTE-----EQGGGGAEGGVVVEGTGESLPPREKLAEPQEVQEAPEBELMKSRE 434
Db 1502 TTLEKTETMTMETHESESRSTSVKKEKT-----PEKVDKPKSPPTKKDKSPEKSITEE 1556
QY 455 MCVSGGDHTQLTDLSPREKTLPHKPEGIVSEVEMLSQERI-KVQGSPLKKLFSSSLGKK 513
Db 1557 I-----KSPVKK--EKSPEKVEKPPASPTKKEKSPKPPASPTKK---SENEVK 1599
QY 514 LSKKKQKRGKGGGDE--EPGEYQHIHTESPESAD-----EQKGESSAS----- 555
Db 1600 SPTKKEKSPKSVVEELKSPKE-----KSPKADDPKSPPTKKEKSPKESKATEDVKSP 1653
QY 556 ----SPEEPE-----TTCLEKPLEAPQDGEABEGTSDGCKREKRGITPWASFKQWTPK 607
Db 1654 KKEKSPKVEKEKPTSPPTKKESSPTTKTDD--EVKSPPTKKEKSPQTVVEKPPASPTKKEKSP 1712
QY 608 K----RVRRSESDKBELEKVKSAT---LSSTDSTVSEMQDEVKTVGEQKPEEPKRRV 660
Db 1713 KSVVEVKSPKESPEKAEKPKSPPTKKEKSPKSAAEVVKSPPTKKEKSPKESKAEKPKS 1772
QY 661 DTSVSMELICVSGSSKRAKASSSDDGGPRTLGDGSHRAEASKDKAGTDAVPASTQ 720
Db 1773 PT-----KKESSPVKMADE-----VKSPTKKEKSPK---VEEKPASPT 1809
QY 721 EQDAQGSSSPAGSPSEGEVSTWESFKRLVTPRKKSXKLEEKABDSVQOLSTEIE 780
Db 1810 KKEKTPKSAAEELKSPTKKE-----KSPSPPTKKTGDESKESPE-----K 1851
QY 781 PSREESNVSIKFTIPGRKKRADGKQOATVEDSGRVEINEDDPNPVAV--VPLSEYNV 838
Db 1852 PEKPKSPPTKPKSPPGPKK-----KKSPEAEKPPAPKLTDLQGTVNKNTDLAHFEV 1907
QY 839 EREKMEAQGNTELPQLIGA--VTVSEB-----LSKTLVHTVSVNAVIDGTRAVTSVEE 888
Db 1908 VEHAETCKWFLDGEITTAQGVTVSKDDQDFEFRCSIDTTMFGSGTVSVV--ASNAAGSVET 1966
QY 889 RSPSWISASVTEPLEHTAGAMPVE-----EVTEKDIIAETPVLTT-----QTLPEKQDA 939
Db 1967 KTELKVLTEPKTKPKBFTDKLRDMETVTKGDTVQMDVIALHSPLYKMYQNGNLLEDGKNG 2026
QY 940 HDDMVTSEVDFTSEAVTATETSEALRTEEVTEASGAETITDMVSQVSLTSDPTTEAT 999
Db 2027 --VTIKNEENKSLIIPNAQDSKGITVEASNEVGSSSS-----AQLTVNPPST---T 2074
QY 1000 PVOEVESGLVDTBEERQTOAILQAVADKVEESOVPAOTVORTGSKALEKVEEVEEDS 1059
Db 2075 PI--VVDGPKSVTIKETETAEFPKATI-----SGFPA-PTVKTWIN-----EKIVEES 2118
QY 1060 EVLASEKEKDV----MPKGPVQEAGAEHLAQGSETGOATPESLEVPVETADVDHATCQV 1115
```

RESULT 9

T15348

hypothetical protein B0350.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15348

R:Gattung, S.

submitted to the EMBL Data Library, February 1996

A:Description: The sequence of C. elegans cosmid B0350.

A:Reference number: Z18332

A:Accession: T15348

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5170 <GAT>

A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208877; PIDN:AAA93447.1; CESP:B0350.1

C:Genetics:

A:Gene: CESP:B0350.1

A:Introns: 48/1; 5039/3; 5116/3

Query Match

Best Local Similarity 6.0%; Score 488; DB 2; Length 5170;

Matches 391; Conservative 249; Mismatches 678; Indels 634; Gaps 89;

```
QY 5 SSTEQRSPQAGSDTPSELVLSGHGPAEASGAAGD----PADADPATKLPQKNGQLSS 60
Db 818 AESERQVPSPVSHQYEPHVETTTTTNTVTSNIYDDEDNVPSSEDPATQHQQ----S 872
QY 61 VNGVABQGDVHVQENQGEQEEVVD-----EDVQRESEEDVREKD 101
Db 873 ETSVHRSHDPDSVEESDGEGLGSKVLGFAKKAGVAGVVAAPVALAAVGAAYADAFED 932
QY 102 RVEEMAANSTAVEDITKQGEETSEIIQIPASENNVEEMVQPAESQANDVGPK----- 155
Db 933 -----DEDDTSHSPES-PVPEYO-SEYOQDDSAQSSTHDFEHMPES 973
QY 156 -----KVFPKVPKFTVKDKNEK-----SDTVOLLTVKK---DEGEAEASVGAGDH 200
Db 974 PIEHEKETBEFDHSHSPESVLSKEKREHQVTSETTTTTTITVREYNDEPDEQ-----EH 1027
QY 201 QEPS-----VETAVGESAKESSELKOSTEKQEGTLKQESSTPIQOAES 245
```

```

Db 1028 GCPHSPASSTHEAHPHIVTTTTVTTRFQ----EEPELEFYKQEDNSKRKSPSSHSQ 1083
Qy 246 DQAAEE-----FAKDEGEKQKEPTKSPESPSPVNSE----- 279
Db 1084 ENLVETTTTTVTSEYDEPEHFEHQFGKSPSPSSHVSESRQVLSPVADPRHVAE 1143
Qy 280 --TTSFKFPPFHGAMGAKTKTSFKSKFE--DLETFEKKKEQGAENKVEEKKTEPPAE 336
Db 1144 TTTTTLVTTRFQ--HDESEKSDSPNRDVVAQSTHSSNIESHQFSEKDDSQGSPFVKS 1202
Qy 337 EOPAE-----DTDOARLSADYEKELPLE----- 361
Db 1203 EDEPVKHSYKRETSTTETTRERFDRPDLSEBRLS--EPAQSPHEVSPITTEENIVAPSS 1259
Qy 362 ----DQVDELAASSEKCAPLAT--EVFEKMEAHQEVAAHVAVSTVEKTEBQGGGE- 414
Db 1260 VKSEYSGEGHVPSVJETKTTTTRREFYDDODEHENQTOSEELRASSIPTEEHEHGHSL 1319
Qy 415 -----AGGVVVGCTGSLPPEKLAERQVPO-- 441
Db 1320 FKETTTTTVTREFYDEPENVEELQDPQSPAPSSHVSGIHASESP--VAQQCEIPQTR 1377
Qy 442 ---EAPAEELMKSRMVCSGGDHQTLDLSP---EKTLPKAPGIVGEVEMLSQERL 495
Db 1378 EFHEDSPAAQYFHEEY-----ESHVLTEQAPLITQHQHP--SQDESDGEGLS---- 1426
Qy 496 KVGQSPFLKLFFSSG-----LKKLSGKKQKQKGGGDEBGE-----YQHHTESP 542
Db 1427 KVLGFAKKGAMVAGVVAALVALLAAGAKAAYALAKKDDDEEDQERESLRLQERSIDSP 1486
Qy 543 ESAD---EKGESSASSPPEPE---ETCL-----EKGLEAPQDGAEBEGTT 584
Db 1487 HASESQIEEHEHFESESPVPSKHHVETTTTTVTTRFDEHPL--VSQELGEGKSP 1545
Qy 585 SGQEK-----KREGITPMASFMM--VTPKKRRPSESDKEEEL 623
Db 1546 ADBEKLPHVETTTTTVTTRFEDKNDSPVSEKEQERTVSRVETFAEBDEBEH-- 1603
Qy 624 KVKASATLSSTDSVSEMODEVKTVGEQKP--EPPKRRVDTSVSEWALICVSSKKRARA 682
Db 1604 ---HYETTTTTVTKEVIDDSQEMGDDEBKQSPQVET-----TTMTSREX 1650
Qy 683 SSDDGGRITLGGDSHRABEASKDEAGTDV-----PASTQODQAQSSSPPEPAG 735
Db 1651 DNDDE--TRSEAGDSHITETKT-----TVVREFHSGQPEETETDEVE-- 1695
Qy 736 SPSEGGVSTWESFKLMPRKSKSKLEBKADSSVEOLSTIEERSREBSWVSIKKFIP 795
Db 1696 -----LPPKIEEDNVSEBSST--SVSRVFRD--EPHITETTTTT 1735
Qy 796 GRRKRAADQEOATVEDSGPVEINEDPNVAVPLSEYNAREKMEAQNTLPPOLL 855
Db 1736 VIREHNEDEETDDQDAAPISFSQ-----EHQDDDSQASHQDHRSP-- 1780
Qy 856 GAVVYSEELSKTLVHTVSAVIDGTAVTSVERKSPSWISASTPELPLEHTAGAMPVEE 915
Db 1781 -----VESEKSVKHT-----TETTT-----TTVTROLTYD-----DE 1807
Qy 916 VTEKDIABETPLTQTLBKGKANHDMWTSYDFTSEAVTATESEALRTEVEITASGA 975
Db 1808 ABE--IRGESPVATE-----EHEHVSSTKSD--ESQHVPSV 1841
Qy 976 EETTDWVASQTLDSPPD--TTEATPVQEVESGVLDTTEERERQOALIQAVADKVEES 1033
Db 1842 IETTTTTVTREFYDDODELQREDHTQSEBRRSI--PTRETHEDHHLK----- 1890
Qy 1034 QVPATQTVORTGSKALEKVEEVEDEVL--ASEKEKV--MPKGPVQEGAEHLAQGSET 1090
Db 1891 ETTTTLVTRREFDEPENVEKLQDSQFSLSPSSHVSESIIVPESPV--AKQCEIPQTRER 1948
Qy 1091 GQATPESLEVPETAVDVHATQVITKQOLMEQAAVAPESSELTLSSTNGSTPLADSDT 1150
Db 1949 HEDSPAAQYFHE--DEYEHQVPTQAPL--LTEQHQHQPESGEE--SPGEGFGSVLGPAPAKK 2003

Qy 1151 A-----DGTQODETIDSQDSKATAAVQSOYTE----- 1179
Db 2004 AGMVAGVVAAPVLAALVAGAKAAYDALKKEDDEDOEE-----RESLIREBSFDSPHA 2057
Qy 1180 -EATAQKEEPESTLPNNVPAQEBEGEE--PQDVLEPTQOELTAAVAVLAKTEVQGEQ 1235
Db 2058 SEQSQIEKENKHFESFPVPSKHHQSSALPOESVSQIEKE-----SRFNDSES 2107
Qy 1236 EYDMDGKXVKEEQQEVFVHSGPNSQKADVTYDSEVWGVACQEKESTEVOSLSLEGEM 1295
Db 2108 EFGVSDHYTEDDQCELSKPSKSGEAFSQFTSEKE-----QDRSDPINS-----Q 2153
Qy 1296 ETDVKEKEKRETKPEQVSEGEQETAP--EHEGTGKRVLTLDWSSSRGKALG-----S 1348
Db 2154 KEDISQFQNESSPEQVKSQPHDEKPLERQGSYSSGY--SPKSPGSGITTGDEEKA 2209
Qy 1349 LGG-----SPSLPDQDKAGCIEVQOSLDT 1373
Db 2210 LSGVQEPEDRPENFASHEKTEATSDENLPESDKAPASP--VPSSDSN--RVLET 2264
Qy 1374 TVTQTAEAVER-----VETVVISETGESPECVGAHLIPAKS-----SATGGMWTLQ 1421
Db 2265 TTTVTRHEPEPEDSHSVYVESQEVSSG--SP-----VPSEKSVDRVLETTTTTVTRE 2316
Qy 1422 H--ADTVPLGPES---QAESIPITVPAP--ESTLHPDLQCEISAS----- 1461
Db 2317 HFEADEIPIITVSESHDDQAS-----FVPSDEDVHQQIOTTTTITVTKHFVFPDEI 2370
Qy 1462 QRESEEDKPKAGPDADGK-----ESTAIKYLKA--PE-----LDEL 1499
Db 2371 DSEHNMESDKTASGSPVPSBEDSGRVETTTTTVTTRHFEPEDDHSPVQTOEYSASBS 2430
Qy 1500 ESKSNKIYLVNIQTAVDQFARTETAPETHAYD-----SQTPVACRLDSREPNRC 1549
Db 2431 PVPSEKSVARVETTT-----TTVTTRHFEDEHILQGGQSGSDQISESTSNMR- 2484
Qy 1550 WTKMDAKKHPVPOFPREDLOVLTYLEMAQP 1581
Db 2485 -----ETSSSPVOGNRDEEFLPALAIAPYKOP 2510

RESULT 10
T26216
Hypothetical protein W06A7.3c - Caenorhabditis elegans
C1:Species: Caenorhabditis elegans
C1:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C1:Accession: T26216
R1:Ainscough, R.
Submitted to the EMBL Data Library, August 1996
A1:Reference number: Z20173
A1:Accession: T26216
A1:Status: preliminary; translated from GB/EMBL/DDBJ
A1:Molecule type: DNA
A1:Residues: 1-2484 <M1>
A1:Cross-references: EMBL:Z78066, PTDN:CA851467.1, GSPDB:GNO0023, CESP:W06A7.3c
A1:Experimental source: Clone W06A7
C1:Genetics:
A1:Gene: CESP:W06A7.3c
A1:Map position: 5
A1:introns: 1827/1, 1866/3, 1963/3, 1990/3, 2262/1, 2289/1, 2339/2, 2463/2

Query Match 6.0%, Score 485, DB 2, Length 2484,
Best Local Similarity 20.2%, Pred. No. 5, 8e-10,
Matches 362, Conservative 244, Mismatches 648, Indels 538, Gaps 69,

Qy 191 ABAASYG--AGNH--QEPVETRVGESAS--KSELKOSTEKOEGTLKQOSSTETPLQAE 244
Db 3 ABAVIGLEDSGNNLRPALLETASDQESSENKEPDASENVETEPEV-----SSAKVIALBS 58
Qy 245 SD-----QAEEBAVDEGEKQKEPTKSPSPSPVNSETTSPFKKFTTHWAGWR 296
Db 59 GDCERGNIKISANENSVEPDGADKPAH--TEQELPTIILPSETIVT----- 102

```

Qy	297	KKTSFKKSKBDDL-ETAEKKKQOEAKVDE-----EBKEKTEPASBEEQEPAPDID-QAR	348
Db	103	-----QLEDHLSETEERKQNEVSSIKSEKNVIGLEDSGDTFRDTSGLAPAEADNEAET	155
Qy	349	LSADYKEKVELPLEDOVDGLEASSEKCAPLATEV--FDEKMEAH-----	390
Db	156	ITTDIF-----VPLE-SAGDIPSNIEKEVASAPDVVGLSEYIIGNIPNAPVVDNDIPNVFT	210
Qy	391	QEVVAEVHVSTVEKTEBEEGGGGEAGGVVV--EGTGESL-----PPEKLAEPQE-VPQEA	443
Db	211	PEVANDETVETFSVTAE-----EASIPVVVELEPIGDEYEFQRPVENFSEPSDNINLEE	264
Qy	444	EPABELMKSREMCSVSGDHTQTLTDLSPSEKTLPKHPIGEGIVSEVEM-----LSS	491
Db	265	SGAEQVLLLENM-----FTPLDELEPOQKILNEKAHI--EIEASGDDEFIKDHPFFVEN	316
Qy	492	QERIKVQGSPLKLFSSSGLKKLGKKQKGRGGGDEEPGEYQHHTHESPE-----SA	545
Db	317	EQNVQEGSRVETVHSFIGLES-SG---VLTGAVSDSVANNVKE-NTESPDIISLEASG	371
Qy	546	D-----EQKGESSASSPEEPETTCI-----EKGP	570
Db	372	DELSKLVEAREIITESKDAYSTDVPESRKTVSDVIGLEAGDKIVSNVSNVMGNPDSEQ	431
Qy	571	LEAPODGEAREEGTTSOGEKKREGITPWASFKNWTPKRVRRPSESDE-----BELEKVK	626
Db	432	LALFQADNVPELPINSEOBETVAVKATESCDHVDPDSOKTLERASSLEBDIMSPVLGITS	491
Qy	627	SATLSSTSDTVSMQDQEVKTVGEEQKPEPKRRVDTSVSWEALICVSGSKKRAKASGSSD	686
Db	492	SQTLSDVLPVISEDQDSIPV-----PVTVEZE-----TSEKLVKASSLEE	531
Qy	687	DGGPRTLGGDSHRAEASAKDEAGTDAVPASTQBQDAQSSSPGAPSGSEGEVSTW	746
Db	532	DVVSPEVLELDD---RVQNKNPSEVTAVDASKTEGD---FSDSPDSRAT-----	575
Qy	747	ESP-KRLVTPR-----KKSXSKLEEKAEDUSSVEQ-----LSTEIERSRE-----	784
Db	576	ETFMELKLVVTENLNPAGDKLSERIQEIARENETISQPKGEDDLENANDPDDETIVEKI	635
Qy	785	-----ESWVSIXKFI-----PQRKRKADGKQEQATVEDSGPVEINDDPNVAVPVLPS	833
Db	636	VSMAESLPIEAVISTEDGTSQDPAQNAIPIUSEETTVDSDQTEEIFTDN-----VKKS	690
Qy	834	EYNAVERKMEAGNTELPLLQAGVTVSSELSKTLVHTVSVAVIDGTRAVTTSVEERSPSW	893
Db	691	KENT-----PKAENDITEINLPGGEGEPD-----NAEKRRNEAVSPND	728
Qy	894	ISASVTEPLEHTAGEAMP-----VEEVTEKDIIASETPVLTQTLPEGKDAHDDM	943
Db	729	ETSEIKODLENLENGASGPNNVQVDAAEQDPTQPTVDETTSKISENMPKAPDTEDDN	788
Qy	944	VTSEVDPFTSAVTATETSEAL-----RTEEV	969
Db	789	ATBIE-----AGLETFEAFGAHVSYLDANIEKLVAMADEPLPVDELVSIEERPEEV	842
Qy	970	TEASGABETDMVSAVSQLTSDPPTTEETATPVQEVESGVLD-----TEEBERQQTALIQ	1024
Db	843	APABSTGEDEDI FRDRDRVTSLTGTGDQNAPIQVIFVSGDGNPDANADQERTSEHNELI	902
Qy	1025	VADVKVESQVPATOTVORTGSKALKKEVEEEDSE-----VL	1062
Db	903	ESDKESSEATIKNEEDVDQPTIQSEBPLTSQGESISSGNKIVANVGVSVLLGAVIPGVCL	962
Qy	1063	ASEKEKQVMPKGPVQEAAGEHLAQSGSETQCATPESLEVPEVTADVDHV--ATCQVIKLQ	1120
Db	963	ASNENEDAHADREVEET-----GDSTRDRPEEETFVYSKLTSMVENILPSTNDENPEAV	1015
Qy	1121	LMEQAVAPASESETLTDG-ETNGTGPLAD-----SDTADGTQOQDETI-----	1160
Db	1016	SMVENLVPVNTGELDSKEDNPDPATAEASHSGEKNLRNDKTTDTRERDPBETITINKLVE	1075

Qy 1161 -----DSQSKATAA VRQSQTVEEBAATQAKE-----PS 1190
Db 1076 NALPTGVGTSETSAPDAQELDVTVDHAGNDTSEYDEAPEKSAGGTVIEKFTSMIES 1135
Qy 1191 TLPNNVPAQEHEGEB-----PGRDLVLEPTOELTAATAAVPLAKTEVG-----QEGEVDWL 1240
Db 1136 ILPVQAPTQPENASHIDQETGGASEIKDDQNQPEEPSAEHQCKFEVSAPPDQESAEPQL 1195
Qy 1241 DGEKVKEQQVFVHSGPNASKAAD-----VTYDSVMGVAGCQEKESTESYOSLSLEEG 1293
Db 1196 EAKKDQDKETI-----ENSEDAKKETVMKLVLSENILPVEAVLPDSST-VTKNSDKK 1249
Qy 1294 EMETTVEKEKETK-----PEOVSEGEEOETAPE-----HECTYGKPVLTLDMPSSERCK 1344
Db 1250 ELETO-ELUSKEIKTGGOPEYVPETSEAFVSDPEIFORVKRASSTEP-----KTQKTEPHA 1304
Qy 1345 ALGLSGSPSLPDQKACICIEVQVQSLDTTVTQAAEAQKVIETVVISSETGESPECVGAH 1404
Db 1305 PIFIQGSTEDEDQSIAANVIDELVHEDD-----EKKYPEVTANISVSAENIDDTTAN 1358
Qy 1405 LLPARKSATGGHWTLQHAEDTVPLGPSQAESIPITVPPESTLHPDLQGEISASQR- 1463
Db 1359 AVP-KTEVSESQLQVATVEFELESAPBEASAIIPEVOEPLEKVEVQPDLSQNSAPHKI 1416
Qy 1464 -----RSESEDPDADGDGKESTAIE----- 1487
Db 1417 IDLHNPIPKDHEDYGNVYPFGTESSESQADGNQENOEEDVVAELNHFPIROWRDD 1476
Qy 1498 -----KVLKAEPBILELESKNKIVLMV 1510
Db 1477 VISLQSLSLVAEVCITDVDASDVNEODEBSTLKILKVPSEPSLLELDFNDPKVIHV 1536
Qy 1511 -----IQAVDQFARTETAPTTHAYDSOTQVPACKLDRRENRCWT 1551
Db 1537 PIPLMEPATMYLEEMVEMIIADVKEVSEMENVTEISEIMAPQVSESTCPPEP----- 1591
Qy 1552 KWKDAKM-----KHPPVQPREDLOVL-TVLEAWAQPRKCLPRLQLKAPVSK 1596
Db 1592 -LADULKVDEDEKTPPEPVPVPGQRIPIEVEQAPTIFQRPRAPKSE 1642

RESULT 11
T26215
hypothetical protein W06A7.3a - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T26215
R/Ainscough, R.
submitted to the EMBL Data Library, August 1996
A/Reference number: Z20173
A/Accession: T26215
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-2607 <MIL>
A/Cross-references: EMBL:Z78066; PIDN:CAB01522.2; GSPDB:GN00023; CESP:W06A7.3a
A/Experimental source: clone W06A7
C/Genetics:
A/Gene: CESP:W06A7.3a
A/Map position: 5
A/Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2412/1; 2462/2; 2586/2

Query Match 6.0%; Score 485; DB 2; Length 2607;
Best Local Similarity 20.2%; Pred No. 6.1e-10;
Matches 362; Conservative 244; Mismatches 648; Indels 538; Gaps 69;

Qy 191 ABASYG----AGDH-QEPPSVETAVGESAS--KSELKQSTEQEGLTKQBSSTEIPLOAE 244
Db 3 APAVIGLEDGGINLRPALETASDQESSENKEPDASENVETEVP-----SSAKVIALESS 58
Qy 245 SD-----QAAEEAKDEGEKEKEPTKGPSPSPSPVNSETTSSFKFFTHGWAGR 296
Db 59 GDCEBGNIKISANENSVPDCAKDKAH---TEQEIPTIILPETTT----- 102

```

QY 297 KTSFKSKEDL-ETAEKKEQEAKEVD-----EKEKTEPASEBEPADTD-QAR 348
Db 103 -----OLEHLSSTEEKQNEVESIKSEKNVIGLESDGDFRRTSGLAEDNEAEAT 155
QY 349 LSADYKVELPDEVDQGLDEASSEKCAPLATEV--FDEKMEAH----- 390
Db 156 ITTDF---VPLE-SAGDIPSENEIKEVASAPDVGLEIITIGNIPNAVNDIDIPNVT 210
QY 391 QEVNAEVHSTVEKTEBEEGGGGEAGGVV--EGTGESL-----PEKLAEPQE-VPOEA 443
Db 211 PEVANDETVETSVTAE-----BASIPVVELEPIGDEYEFORPVENFSPSDNINLEE 264
QY 444 EBAEELMSREKCVSGGDHTQULTDSPEKRTPKHPEGVSEVEM-----LSS 491
Db 265 SGAEQVLENNM-----FTPLDELEPOQKLINEKAEHI--EIASGDPEFIKHDPPEVN 316
QY 492 QERIVQSGPLKRLPSSSGLKKLSGKKQKKGKGGGDEPPEYOHHTESP-----SA 545
Db 317 EGNVNEQSRVETVHSFIGLES--SG--VGLTGAVSDSVANNVKE-NTESPDIISLEASG 371
QY 546 D-----BOKESSASSPPEETTCI-----EKGP 570
Db 372 DELSLVAREIITTESKAVSTDPESKTVSDVIGLEBAGPKIVSNNVSNVGNPDESQ 431
QY 571 LEAPDGEAEGETTSDGEKKREGITPMASFKKMVTPKKVRKRPSESDKE---ELEKVK 626
Db 432 LALBQADNVPELPINSEQETVAVAKTESCDHVVDSCQTLERASSLEEDIMSPEVIGITS 491
QY 627 SATLSTSTVSEMODEVTVGEBQKPEPKRVPDTSVEMWALICVSSKKARAKASSD 686
Db 492 SQTLDYLVFVISEDQSTIV-----PVTEVE-----ISEKVKASSLEE 531
QY 687 DEGPRTLGDSHRAEASKDEAGTDAVPASTOEOQAQSSSPPEPAGSPBEGVSTW 746
Db 532 DVVSEVLELD--RVQKNPSEVTLAVDASKTEGD--FSDSPDSRAT----- 575
QY 747 ESF-KRLVTPR-----KSKSKLIEKAEDSSVEQ-----LSTLEPSRE----- 784
Db 576 ETFMEKLVVTENLLPAGDKLSEERIQTRENETTISQPKEDDLENNANPDETIVEKI 635
QY 785 ---ESWVSIKFFI-----PGRKRKADGKQEOATVEGSPVEINEDPNVPAVPLS 833
Db 636 VSMASSLTEAVISTEDGCTSDQPAQNAIPSEETTVDSQTEELFTDN-----VKKS 690
QY 834 EYNAVEREKMEAQNTLPGALVAVVSELSKTLVHTVSAVITDTRAVTSEERSPSW 893
Db 691 KENT-----PKAENDTEINVLPGEGEGRPD-----NAKKNBAVSPND 728
QY 894 ISASVTEPLEHTAGEAMP-----VEVTEKOTIAETPVLTQTLPECKAHADM 943
Db 729 ETSEIKQDLENLNGASGPDNNVQVDEAAQEDPTDETVETTSKISENMPKAPDTEEDN 788
QY 944 VNSVDFTSEATVATETSEAL-----RTEEV 969
Db 789 ATEPIE-----AGLETTAFGDABEVSYLDANIEKLVMADEPLVYDELVSTEERPEV 842
QY 970 TEASGABETTDWVSAVSQULTSPDTEATPVQEVESVGLD-----TEEBERQTAILOA 1024
Db 843 AAASETGEDEDFRRDRRIVSLTGTDQNAPIQVIFVGGDENPDANAAQERISHNELI 902
QY 1025 VADKVESQVPATQTVQRTGSKALEKEVEVEBDS-----VL 1062
Db 903 ESDKSESEEAITKNKEEDVDQPIQSEEPITLSQEGESSIGNKIVAVVGSVLLGAVIPYVL 962
QY 1063 ASEKEDVMPKRPVQAGAHLAQSGSETGOATPESLVEPVAVADVHV--ATQVVIKLOQ 1120
Db 963 ASNENEDAHADREVEET-----GDSTRDPEETFEVSKLISWENILPSTINDENPEAV 1015
QY 1121 LMEQAVAPESSETLTDG--ETNGSTPLAD-----SPTAGTQODETI----- 1160
Db 1016 SKVENVLVNTGGLBESKSDNPDAPFAEHAHSGEKNLRANDKTDTTRDRDEBEETILKLVE 1075
QY 1161 -----DSQDSKATAVAHQVQVTEEEAATAQKE-----PS 1190

```

```

Db 1076 NALPVTGTSFTTEVSAPDAQELDETVDVHAQNDTSEVEDAPEKSAGTVIEKTSMTES 1135
QY 1191 TLPNVNPQOEHBGE-----PGRDVLEPTQOELTAAVPAVLAKEVNG-----OEGEVNWL 1240
Db 1136 ILPVQAPTPQPNASHHIDQETGASSEIKDQNPPEEFSAHQKFEVASBPDDQSABEQ 1195
QY 1241 DGEKVEEQEVFVHSGPNQQAAD-----VTVDEVMGVAGCOEKSETEVQSLSEEG 1293
Db 1196 BAKDQDKETI-----ENSDPAKRTVMEKLVSVENITLPEAVLPSDST-VTNKSEBCK 1249
QY 1294 EMENTVEKEKETK-----PEOVSEGEQETAAP-----HEGTQKPLTLTDMPSSEBCK 1344
Db 1250 ELETQ-ELSSKEIKTSQGPETVPETSEAFVSDPEIFQVKAASSTEP---KTQKTEPHA 1304
QY 1345 ALGSLGSPSLPDQKACIEVOVQSLDTTQTAEAVERKIVTVVISEGESPEVCAGH 1404
Db 1305 PIFIQGSTEDDEQSIANVIDELVHED-----EKKVPEVTANISVASINIDSTTAN 1356
QY 1405 LIPAEKSSATGGHWTLOAEDTVPLGPESQAESIPITVTPAPESTLAPDLOGEISASOR 1463
Db 1359 AVP--KTEVSSBQLOVATVEPELSEAPESBAIPEVQEPLEKVEVQPDLSQNSPAPKI 1416
QY 1464 -----ERSEEDKPDAGPADGKESTAI----- 1487
Db 1417 IDLHFNITKDHEDYNDVYPPGTSESSBQADGNQOEEDVVAELNHPRIQWRDED 1476
QY 1488 -----KVLKAEPEILEKSKNKIVLNV 1510
Db 1477 VISLSKSLVAEVCITDVDAVDNQBDEESTLKILKVSEPSLELDTNIDPKVIVH 1536
QY 1511 -----ICTAVDQFARTTAPETAHVSOTQVYACRLDSREPNRCWT 1551
Db 1537 PIPLEPATMYLENNVEMWIIDAVAEVSEMEVTESEISMAPOVSESTCPIPEP----- 1591
QY 1552 KKKAKM-----KHPVQPREDDQVL-TVLBAMQPKKCLPRLQKAPVK 1596
Db 1592 -LADIKLVEDEKTPPEPEVVPQGVQERIIPIEVEQAPTIIPORPPKPKSE 1642

RESULT 12
S00485
gene 11-1 protein precursor - malaria parasite (Plasmodium falciparum) (fragments)
C1Species: Plasmodium falciparum
C1Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000
C1Accession: S00485
R1Schierf, A.; Hildich, C.; Sieg, K.; Matrei, D.; Mercereau-Pujalon, O.; Mueller-Hill, B
EMBO J. 7, 1129-1137, 1988
A1Title: The 11-1 gene of Plasmodium falciparum codes for distinct fast evolving repeats
A1Reference number: S00485; MUID:88296416; PMID:2841111
A1Accession: S00485
A1Molecule type: DNA
A1Residues: 1-1315;1316-1485;1486-1657;1658-1729;1730-1948 <SCH>
A1Cross-references: EMBL:X07453
C1Comment: This protein is associated with the membrane of red blood cells at the schizont
A1Gene: 11-1
A1Introns: 71/3
C1Keywords: tandem repeat
F1-71/Domain: signal sequence #status predicted <SIG>
F172-1948/Product: gene 11-1 protein (fragments) #status predicted <MAT>

Query Match 6.0%, Score 483.5; DB 2; Length 1948;
Best Local Similarity 21.9%, Pred.No.4,9e-10;
Matches 371; Conservative 252; Mismatches 619; Indels 455; Gaps 76;

QY 28 GHGPAEASGAAGPADADPATKLPQKNQULSVNGVABQGDVHVQENQSQEVEVDE 87
Db 283 GNEGEGDGHGSGAGDAGD-----RGDRNKQDKDNEKNKKGNKDDDEEBEED 335
QY 88 DVGQGESDVAKQVEMMANSTAVEDITKDGQETSEI-----IEQIPAEENV 139
Db 336 DEKKEKDKQKDEKDEKDEKDEKDEKDEKDEKDEKDEKDEKDEKDEKDEKDEKDEK 395

```

Qy	140	EMVQPABSQANDVGFKVPKVFVGFK-----FTVKDKNEKSDTVQLLTVKKD	186
Db	396	KIERDK-----EFEGFKDEYFTOMEGGAYGITVEHQGLRILT-----	436
Qy	187	EGEGAEASVAGDHQBPS-----VBTA VGESAKESEL-----	222
Db	437	--EGVEGYEKVIEMEQTKDLBFLPTNCKGSDDFVSEMEERLFNFVS YMEEVVKDMF	494
Qy	223	TEKQOGLTQ-----EOSSTEIPLQAESDQAABEEAKDEGEOKOEKPTKSP	269
Db	495	IEREQGTEEBVIHEDLSDEKIGIEKKELETFVQDVSDVLQDEQLIEBKUTEKLTTEE	554
Qy	270	ESPSSPVNSTTSSPKFFTHGWAGRKTSPKSKEDDLTAETAKRKQEAKVEDEEK--	328
Db	555	EKUTEE-----EKLTEBEKTEDE-QLIIEEKLTEBEKLTETEKL	593
Qy	329	EKTPASEEOEPAD---TDQARLSADYEKVELPLEDOVGOLEASSEBKCAPLATEVPDE	385
Db	594	TEBEKLTEBEKLTETEKLTEEBIFIB---EDKI-----SPKEVEPL-EDVIME	644
Qy	386	KMEAHQEVVAHVSVTEKTEBEBOGGGBAEGGVVGTGSGSLPPEKLAE-----	435
Db	645	KMEKEYKEILHEGKSVVVTQMDEKLLSENLEDOLPIEBIIBELVDEELVEEVTAEEIVEE	704
Qy	436	-----POEVQEAEPABELMKSRMCVSGGDHTQLTDLSPBEKTLPHKEGIVS	484
Db	705	FVVEVEITEKIPKELVEENVPEED-----VDVVPES--LVEEPDESVE	746
Qy	485	EVENLSSQERIKVQGSPLKLGSSGLKKLKQKGRGGGOEPE-QYHIHTESPE	543
Db	747	EVITTEELTEEV-IPBELIEKVVEEELVEKVAQKIV-----EVIPEELIEEVIETKP-	797
Qy	544	SAREKQGE--SSASSPDE-----PETTCLKGPLEAPQDGAEBEGTTSOGCKREGITP	596
Db	798	-ABEVIPELVEEVIPEIBEIVEELPPEWV-KYVIPSEWVEEVIPEBVSEVPEEIVSEWVP	855
Qy	597	WASFPMVTPKRVRPSPSEDKEELEKVKYSATLASTDSTVSEMQDEVKTYGEEKPEEP	656
Db	856	-----SELV-----EEMKPEVEFEVISSEL-----VEEMKPE--EVVEEVTPE--	892
Qy	657	KRRVDTSVSWEALICYGSSKKRKARKASSDDBGGPRTLGGDSHRAEASKDKAAGTDAVP	716
Db	893	--KVSEWVPEELV-----EEMKPEEUVVEEVP	918
Qy	717	ASTQEQ---DQAGSSPPAGSPSCEGVSWTWSFKLVTPRK-----SKSKLEBAE	768
Db	919	EELVEEVIPEEVEEVIPEEWHEAKHE-----ELFEKL-VPEVEBEIIPKELVVIPE	972
Qy	769	DSSVEQLSTIEPSR--EESWSV---IKKFIPGRKKRA---DGKQE---QATVEDSGVP	817
Db	973	EVLPEELIEEVIPEEIEEVIPEEVEEVIPEEIQVEVIPEDLMEEVPEEIEEKIPE	1032
Qy	818	EINBEDPN---VPAVPLSEYNAREKMAQNTLELPOLLGAVVSVBELSKTLV-----	869
Db	1033	EIVEERIETHESIIVEIIP---BEWVEEVPFAEVEEMIPEKIVEEVIPEELVEEVPVEVL	1088
Qy	870	-HTVSVAVIDGTRAVTSVEERSPSWISASVTPLEHTAGEAMP--VEEVTEKOIIAET	926
Db	1089	EEVVPVLESEVIPEEVEEVIPEVLEEVIP--BEWVEEVIPEEFIEEMIPEEIVSEV	1146
Qy	927	P-VLTOTLPKGADHDMTVSEVDFTSATATATSEALRTBEVTEASGETTDMVS AV	985
Db	1147	PEVVEVIPE--ELVEEVPVEL---LEEVISELLEEVIPEELVEEVPPELVPEEMKPE	1201
Qy	986	SQ LDT-SPDT-TREATP---VOVESGLVLTDEE-----BERQ	1017
Db	1202	EVLSEVIPEVLSEVIPEELVBEVIPLEVSESEVEEVLVEEVIPELLVEEVPVPEEVEVQ	1261
Qy	1018	TQAILQAVADKYEE-----SOVPATQTQORTGSKALEKV--EEVEED-----S	1059
Db	1262	EQLVEEFPPEEVEEVLVPEELVEEVIPEEVEEVEEVEEVEEVIIPPEEIVVRPGPEELVE	1321

[illegible]

```

Db 527 KAATSKPVTYDK--VVKKE-----IKTKPEKKEKKEKVAKAKEDTPLKDKDK 576
Qy 272 PSSPVNSETTSFKKFTTHGAWGRKRTSPKSKEDDLETAERKKEQAEKVDEEK--E 329
Db 577 PK---KEEAKKEIKK-----EIKKEEKKELK-KEVKKETPLKDAKKEVKDE 619
Qy 330 KTEPASEEOPEDBDQAALADYKVELPLEDQVODLEASSEKCAPLATEVPEDEKMA 389
Db 620 KKEVKKEEKEPKKEI--KKISKDIKK-STPLSD-----TKPKAALKPKYAKKEEPT 667
Qy 390 HOEVAEAVSTVEKTEEEOGGGGEAEVVEGTESLPPREKLAEPQAEPAEAEEL 449
Db 668 KKEPILAGLKKKGKVKVTKKEGKTEAATAVGTAAVAAAAGV-----ASGAKEL 720
Qy 450 MKSREMGSGDHTQ-LTDLSPREKTLPRHPEGIVSEVMSSQERIKVQSGPLKFLSS 508
Db 721 BAKRSIMSPEDULTKDFELKABEIDVAK---DIKQLELIEDEKCLK-----765
Qy 509 SGLKILSGKKQKGGGGDEBERGEYQHTHSPESADOKGBSSASSPEBPEETCLEK 568
Db 766 -----ETEPGEAYVIQKE-----TEVERK 783
Qy 569 GPLBAPOGEAEEGTTSDEKKRKEGTTPMASFKQWTPKKAVRPRSESDKEELEKVKSA 628
Db 784 GSAESPDEGI---TTTEGECE-----QTP-----EELERV---812
Qy 629 TLSTSTSVSEMODEVKTVEBOKPEPRKRVDTSVSWEALICVSSKKRARKASSDDE 688
Db 813 -----EKQGVNDIEKFEDGAGFEESBAGDYEEKA-----843
Qy 689 GGPRTLGSHHAEPASKKEAGTAVPASTQEQDAQSSSPERPGSSBEGVSTWES 748
Db 844 -----ETEEKEEPEEDG-----EDNVSSASAH---SPTEBEIKAKAA 879
Qy 749 FKRLVTPRKSKSLKEEKEDSSVEQLST-EIEPSREESWVSIKFIPRRKKRADGKE 807
Db 880 DVH-IKEKESVAGSDRAREEDMDLEKGEAQSEEG-----EEDDKAE 925
Qy 808 QATVEDSGVEINEDDPNVPVAVPLSEYNAVEREKMAOGNTELPOLLGAVVSEELSKT 867
Db 926 DAREEDHEPDTEARD-YMAVAVDKAABAGVTEQYDFLGT---PAKQGVGSPSREPAS 981
Qy 868 LVHTVSAVIDGTRAVTSVEER---SPSISAS--VTPELEHTAGAMPVEVEVEKDI 922
Db 982 S1HDETLPGSSSEATASDENREDDPEEFATISGYTOSTIEBISSEPTMDEKSTPRDVM 1041
Qy 923 AERT-----PVL--QTLPEKG--DAHDMVTSEVD 949
Db 1042 TDETNNETTESPOEFVNITKTESSISYQFYSKFPVASFNGLSGSKTDATDGR-----D 1096
Qy 950 FTSEAVTATETS-----EALRTEVTEASGAETTDWVS-----AVSOLT 989
Db 1097 YNASSSTISPPSMEEDKFSKALRDAYRPEETDVYTGAEIDKIDVSDRLSPAKSPSL 1156
Qy 990 DSDPTTEATPVQVEESGLTDEBERQTOALLQAVADKYKESQVPAQTOYQRTSKAL 1049
Db 1157 PPSPPSIEKTPIGERSVNSLTPE-----IKASABEBAVAVSPGVTOAV 1202
Qy 1050 -----EKVEVEDESEVLASEKEKOVMPKPVQOAGAEHLAAGSEGTGATPESLEV 1100
Db 1203 VEHKASPEEKTLEVPSSQSTVGSAGHTPYOOSPTEDESSSHPLEVTENAOAVPISPEF 1262
Qy 1101 PEVTADVDHVAIC---QVTKLQOLMEQVAPESSETLTDSETNGSTPLADSDVADGTQ 1156
Db 1263 TEAKDENESSISPMDEVPDSESPLEKVLSPLRSPPLGSESAVDFLSADDKALGRS 1322
Qy 1157 DETIISQDSKATAAAYQSOVTEBEAATAQKEBPSTLPNNVPAQEEGEPGRDVLPTQ 1216
Db 1323 ESPFGKNGKQGFSDKESFVSDLTSLYODKQEKRAIGFIPKIEDPSPPKASDAIMSS 1382
Qy 1217 ELTAAVAVPLAKTEVQEG---EVDMLDGEKVEBOEVFVHSGPNQKA-----ADVT 1266

```

```

Db 1383 QSALA-----LDERKLGADGSPPTQVDVQFSGFKEDTKMSISEGTVSDSKATPVDEGAEDT 1438
Qy 1267 YDEVMVAGCO-----EKSTEV-QSLSEBEG-----1294
Db 1439 Y-SHMEGVAASVATASATSSPEPTDDVSPSLHAEVSGSPHSTEVDLSVSVQTPPTF 1497
Qy 1295 METDVEKERE-----TKPEQVSEBEGQETAPEHGGTGVKPLTILDMSSSEKGLG 1347
Db 1498 QETEMSPKEEPRPMSISPPDSPKTAKSRTPVODHREQSSMSIERFGQSPERSHAMD 1557
Qy 1348 SLGSPSLPDDKAGCI-----EVQVQSLDPTTVQTAANV-----EKVIET 1388
Db 1558 FSRQSPDHPVTG-AGMLHTENGPREVDYSPSDIDQSSLSHKIPTEBPSYTOQNDJSEL 1616
Qy 1389 VVISTGSESPCEVGAHLPAEKSSATGSHWTLQHAEDTVPLGPESQAESIP1IYTPAPES 1448
Db 1617 ISVSQVEASPTSSAH-TPSOIAS-----PLQEDTLSD-----VVPPRDM 1655
Qy 1449 TLHPDLQGE-ISAQRER-----SEBEDKPDAGR-----DADKESTALEKVLKA 1492
Db 1656 SLVYASLASEKVQSLGEKLSPKSDISPLTPRESSFTYSPGFSDTSGAKESTAYOTSS 1715
Qy 1493 EP 1494
Db 1716 PP 1717

```

```

RESULT 14
D96796
probable heat shock protein, 53413-59028 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Dates: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96796
R:Thellogis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
  Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
  ausen, N.F.; Hughes, B.; Huzar, L.
  Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
  C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Matzli,
  Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
  ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96796
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1871 <STO>
A:Cross-references: GB:A005173; NID:96143906; PIDN:AAFO4452.1; GSPDB:GN00141
A:Gene: F28016.15
A:Map position: 1

```

```

Query Match          5.8%; Score 466.5; DB 2; Length 1871;
Best Local Similarity 20.5%; Pred. No. 1.9e-09;
Matches 331; Conservative 248; Mismatches 588; Indels 447; Gaps 68;

```

```

Qy 50 KLPQKNGQLSSVNGVAEQDVHVEENQ-----GQEEVVED 88
Db 310 KLOKEEBQHSK---EQGG--HSKEENKELVEEKTPEAETTRINDILQPGQEIIVPEVD 363
Qy 89 VQGRSEEDREKOR-VBEMANSTAVEDI-TKQGEELSEIIQIAPAE-----N 136
Db 364 TLGKTSIDGKEKONIVKKEIRNGDATEIIDKMGVFASNIADTGMNSDEFSKLEB 423
Qy 137 NVEEVPQAE-----SOANDVGFKRVFKFTVKKDKNKKSDTVOLTLVKDE 187
Db 424 EVDKAVKEKKDROENDKXGAQSEDSITKLOEIGQOQOGK-RDKKENIKELA---E 478
Qy 188 GEGASAVGAGDHQBPVSVETA VGESAKSESELKOSTEK--OEGTLKQBSSTELPQAES 245
Db 479 GQASA-----EKNIKNDILKPVQKRSBGKHKIQTPQEBTKQEGVNEKIM--ET 528

```

```
QY 246 DQAAEEAKDEGEE---KOE-KEPTKSPSPSP--VNSSETSSFKFFTHGHWGRKKT 299
Db 529 GKINEDGTRKVOEMTROELDPASEKENRSRELVKNTNDEKK-----EKBI 579
QY 300 SFKSKEDDLTAETAEKKEQB-AEKVDSE-----EKEKTPASSEEOBPAEDTDQARL 349
Db 580 ACTERKESDRPKILREQEVADEVAEDTKPSIYGEVKEEBEIAKKEKEFGSDDDIARI 639
QY 350 SADYEVK-----ELPLEDQVGD-----LEASSEKCAPLATEVDFDEK 386
Db 640 VRDTEQLDSNAMQGBEKDQMIQELVLEEKVDCGGKGIIVAEATKAENNKSKRVOETEEOQ 699
QY 387 ME-----AHQEVVAIVHVSTVEKTEEBEOGGGEAEAGGVVEGTGESLPPEKLAEPQEV 440
Db 700 LKEDTCGHFKQKLIIEGSDHGEVDVEKGRKTEA-----EKRIK 741
QY 441 QEAPEAEELMKSMCMVSGG--DHTOLTDL-----SPEKTLPKHPGIV---- 483
Db 742 DRAREAEII-KEKDLGVSGRYIKGTIIKELVENRGYIRNEHEKKKDDANRPEKITGTIK 800
QY 484 SEVEMLSQ-----ERIK-----VOGSPCLKLFSSSGLKKLGGKQ--- 519
Db 801 QELVLSNQLRQBNVEDGDKTQLBEVKIKDCBEEGSEBSKIKTDVVRKVQGIKEEEL 860
QY 520 -KKGKGGGDEEPEYQHHTHTEPSADEQKGESSASSPEEPTTCLKEGPLEAPQDGE 578
Db 861 YKPKRHHG-----TKITELVEETGD---YEKQEEKET-----AESDIE 896
QY 579 ABEGTTSS--DGEKKREGITPWASFCKMWTPKRVRRPSSDKKEELEK-VKSATLSTDS 635
Db 897 AECGSURKVDGIEHELHPKIHKEKDNRRNVTGAKPSGOEKEKEKIVESMTTENDN 956
QY 636 TVSEMDEKVTVG-----EEOKEPEPKERVDTSVSWEALICVGS 675
Db 957 SIDVQETKXKPCRLSHDKRYKIQELMEAGHNDKKEEQNVTAVELEET---ERVS 1013
QY 676 KPARKASSDDGG-----PRTLGGSHRA 701
Db 1014 SKKVQSGMEDDNGSGFHEFERKSYEDWTHKREKRKVLVEBEETYPKDKHTGGEDHND 1073
QY 702 EEAASKDEAGTDAVPASTQS-----ODQAQSSSPPEAGSPSEGEVSTWESFK 750
Db 1074 HKEEEOKENVIAKAEINTEEDSPKVEETEKQDHGELKRSWQAKKQETEEKDKTRAMEK 1133
QY 751 RLVTPRKK-----SKSKL-----EKAEDSSVEOL-STEIEPSRE-----ESW 787
Db 1134 NETVRRKQTKDGLKREGDEPELGGHERRGEEDRIEELVETISDHKEVKYKKDEY 1193
QY 788 V----SIKXFIQRRKKRQADKQEQATVEDSGPVEINEDDPNPAVPLSE-----YNA 837
Db 1194 ILRSQDTGKVDLGERERRSKQRKIHKSVDEIDGQDEDAEAEAAVVRNENGSRKVQT 1253
QY 838 VEREKMAQGNTELPOLLGAVVYSELSKTLVHTSVAVIDGTRAVTSVEERSPSMAS 897
Db 1254 IEESEKHKEQNKIPETSNP-EVNEE-----DEERVV-----EKETKVEAH 1294
QY 898 V-----TEPLETAGEMPPVEVTEKDIABETPVLITQTLPEGKADHDMVTSEVDFT 951
Db 1295 VQLEGTENCKDDGEGRR--EERKQKQTAEN--MLRQRF---KTKSDDGIVRKIQET 1347
QY 952 SEAVTATETSEALRTEVTASGABETTDMSAVSOLTSPTDTEBATPVQEVESGLDT 1011
Db 1348 KE-----BEPDEKKSQESSSHVVKLVAE--DG-----SLRNG-LEF 1380
QY 1012 EEEERQTOALQAVADKVEESQVPATQTVQRTGSKALEK-----VEEVEDSEVLAS 1064
Db 1381 SEKESTVSKMLKLDSEKKEEHKIRKPTTEERSNAPVIEKQGNKNAEEMQDKIDRRGK 1440
QY 1065 EKE-KDVMKPGPVQEAAGHAQ--GSETGQATPESLEVPVTDVHVATQCVIKIQ 1120
Db 1441 NQBIKQEPYGVLRNHEHDKITYHKGEEKGTA-----ENVSS---TKIQ 1483
QY 1121 LMEQAVAPESSETLTDSETNGSTPLADSDTADGTQOQDETIDSDQSKATAAVRSQVTEEE 1180
```

```
Db 1484 TXDELEKRPKPSISENNN-----IHFMDSSQDIEEKSGDQAEKAYAKONKIQE-- 1534
QY 1181 AATAKEEPSTLPNNVPAQOEHEGEPGRDVLPTQOELTAAAVPVLAKTEVQO-----E 1234
Db 1535 -----VMNDEKKEEY-----HISERVNEMAKRILQVESKANDGSSKKNETE 1577
QY 1235 G-EVDWLQGEKVKESQEVFVHSGPNSQKAADVYDSEVNGVAGCQEKESTEVQSLSLEG 1293
Db 1578 QGESTGLGRKKRENHQELVELETSDQKKG--VKDEVVGKAEIIEDEYDSSRKIHEHEE 1635
QY 1294 ENETDVEKEKRTKPRQVSEGEQETAAPHEHGTGKPVLTLDMPSSERGGKALGSL--- 1349
Db 1636 RMSDKLEMHGEEMSEKJAEETSDEGAENGRAG-----KSRDDGFGKVRKIE 1686
QY 1350 ----GGSPLPDQDKAGCIEVQVQSLLDTVTQTAABAVEKVIETVVISSETGESPE 1399
Db 1687 VQKNDQSFVEKDTSGRAKENLNDEEPTKTETKATDNESRKIHQIKEQGTSEQ 1740

RESULT 15
T34518
nestin - golden hamster (fragment)
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34518
R:Steinert, P.M.; Chou, Y.H.; Prahlad, V.; Parry, D.A.D.; Marekov, L.; Wu, K.C.; Jang, S
submitted to the EMBL Data Library, December 1998
A:Description: A high molecular weight intermediate filament associated protein in BHK-2
with type III vimentin and type IV a-internexin.
A:Reference number: Z21538
A:Accession: T34518
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1804 <STE>
A:Cross-references: EMBL:AF110498; NID:g4063502; PID:g4063503; PIDN:AAC98312.1
A:Experimental source: cell line BHK-21

Query Match 5.7%; Score 456.5; DB 2; Length 1804;
Best Local Similarity 20.8%; Pred. No. 4.2e-09;
Matches 358; Conservative 249; Mismatches 575; Indels 541; Gaps 79;

QY 2 GAGSSTEQRSPPQ-----AGSDTPELVLSGHGPAABASGAAGDADADATKLPQKN 55
Db 246 GVLTELEAGGQPGHPEDATASAPS---LSPHPVLEAKD--GDSTES----- 290
QY 56 GQLSSVNGAEGDVHVQENQEGQEEVVVDVCGRESEDEVREKDRVVEEMAANSTAVED 115
Db 291 -----RGSSIFQED--EGQIWELV-----EKEAAIELKVESSLAQE 324
QY 116 ITKDGQETSETIEQIPASENNVEMVQPAESQANDVGFKVFKFVGFKFTVKDKNEKS 175
Db 325 TOEDGLH-----TEIIDSQGLQK----- 344
QY 176 DTVQLLVTKDGEAGASVAGADHQEPSEVETAVG-ESASKESELKQSTKEGTLKQSQ 234
Db 345 ETLEAL-----GEEPLMSLKIQNHETPGKNCNLSRSDENQGTILKSPEEEKQTLKSL 398
QY 235 SSTEPLQAESDQAAEEAAKDEGEKQEKPTKSPESPSPVNSSETTSSFKFFTHGWAG 294
Db 399 BEKDQVEVEKTLKGVPELSKPLGKE---DPRIEQELMSPEGTLETLSF-----IG 446
QY 295 WKKTSFKKSDEDDLET--AEKRKEQAEKVDEEKEKTEPASEQEPAPEDTDQARLSAD 352
Db 447 KNEEVRVSSEENIESLAAFKESQHPGLGCEEBEIORVERLIEKE-----QGESLSP 500
QY 353 YEKVELPLEDQVGD--LEASSEKCAPLATEVDFDEKMEAHQEVVAHVSTVEKTEEOG 410
Db 501 EE-----EDQETDRPLEKENGEPKLPVEEB--DQLFETLIEKEGQESLSSPEEEDQET 552
QY 411 GGEAGGVVVVGTGESLPPPEKLAEPQEVQAE-----PABELMKSRMCVS 458
Db 553 RPLEKEEDQLVERLVEKEGQESLSSPEEEDQETDRPLEKENGEPKLPVEEDQLFETLIE 612
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2002, 21:09:06 ; Search time 49 Seconds

(without alignments)

1350.943 Million cell updates/sec

Title: US-09-902-432-4

Perfect score: 8073

Sequence: 1 MGAGSSTQRSPQAGSDT.....AWAQRKCLPRLQLKAPVSK 1596

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4294	53.2	1781	1 AKAC HUMAN	Q02952 homo sapien
2	501	6.2	2464	1 MABP MOUSE	P14873 mus musculus
3	493	6.1	3924	1 ANK2 HUMAN	Q01484 homo sapien
4	492	6.1	2468	1 MABP HUMAN	P46821 homo sapien
5	470	5.8	2459	1 MABP RAT	P15205 rattus norv
6	433.5	5.4	2774	1 MAPA RAT	P34926 rattus norv
7	421	5.2	2805	1 MAPA HUMAN	P78559 homo sapien
8	411.5	5.1	3562	1 PGCV CHICK	Q90953 gallus gall
9	408.5	5.1	4687	1 PLEI RAT	P30427 rattus norv
10	402.5	5.0	4473	1 PLEI CRIGR	Q9J155 cricetus
11	400.5	5.0	4684	1 PLEI HUMAN	Q15149 homo sapien
12	395	4.9	3396	1 PGCV HUMAN	P13611 homo sapien
13	388	4.8	1805	1 NEST RAT	P21263 rattus norv
14	388	4.8	1861	1 MAP2 RAT	P15146 rattus norv
15	385.5	4.8	1087	1 NPH MOUSE	P19246 mus musculus
16	382	4.7	2116	1 MY52 DICDI	P08799 dictyostell
17	379	4.7	2349	1 TPR HUMAN	P12270 homo sapien
18	377.5	4.7	1827	1 MAP2 HUMAN	P11137 homo sapien
19	376	4.7	3256	1 KI67 HUMAN	P46013 homo sapien
20	374	4.6	1233	1 YF16 YEAST	P43597 saccharomyc
21	374	4.6	1637	1 MRSP STRAU	P80544 staphylococ
22	371.5	4.6	1828	1 MAP2 MOUSE	P20357 mus musculus
23	371.5	4.6	3911	1 AKA9 HUMAN	Q99996 h a-kinase
24	371	4.6	1658	1 M67 YEAST	Q03661 saccharomyc
25	370	4.6	1391	1 MS72 DROHY	Q08696 drosophila
26	369.5	4.6	3381	1 PGCV BOVIN	P81282 bos taurus
27	367.5	4.6	1411	1 TCOF HUMAN	Q13428 homo sapien
28	367	4.5	1020	1 NPH HUMAN	P12036 homo sapien
29	366.5	4.5	1972	1 NFI HUMAN	Q12888 homo sapien
30	363	4.5	831	1 NFI RAT	P16884 rattus norv
31	363	4.5	2453	1 NCR1 MOUSE	Q60974 mus musculus
32	354	4.4	1618	1 NEST HUMAN	P48681 homo sapien
33	352.5	4.4	1616	1 P200_MYCGE	Q49429 mycoplasma

RESULT 1

AKAC_HUMAN

ID AKAC_HUMAN STANDARD; PRT; 1781 AA.

AC Q02952; Q99970; Q00498; Q00310;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE A-kinase anchor protein 12 (A-kinase anchor protein 250 kDa) (AKAP

DE 250) (Myasthenia gravis autoantigen gravin).

GN AKAP12 OR AKAP250.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Heart;

RX MEDLINE=97153077; PubMed=9000000;

RA Nauert J.B., Klauk T.M., Langeberg L.K., Scott J.D.;

RT "Gravin, an autoantigen recognized by serum from myasthenia gravis

RT patients, is a kinase scaffold protein.";

RL Curr. Biol. 7:52-62(1997).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Umbilical vein endothelial cells;

RX MEDLINE=98269042; PubMed=9604001;

RA Sato N., Kokame K., Shimokado K., Kato H., Miyata T.;

RT "Changes of gene expression by lysophosphatidylcholine in vascular

RT endothelial cells: 12 up-regulated distinct genes including 5 cell

RT growth-related, 3 thrombosis-related, and 4 others.";

RL J. Biochem. 123:1119-1126(1998).

RN [3]

RP SEQUENCE OF 43-1781 FROM N.A.

RC TISSUE=Umbilical vein endothelial cells;

RA Bowditch R.D., Ginsberg M.H.;

RT "Sequence of gravin cDNA isolated from a human umbilical vein

RT endothelial cell library.";

RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE OF 1477-1781 FROM N.A.

RC TISSUE=Umbilical vein endothelial cells;

RX MEDLINE=92395179; PubMed=1522245;

RA Gordon T., Grove B., Loftus J.C., O'Toole T., McMillan R.,

RL Lindstrom J., Ginsberg M.H.;

RT "Molecular cloning and preliminary characterization of a novel

RT cytoplasmic antigen recognized by myasthenia gravis sera.";

RL J. Clin. Invest. 90:992-999(1992).

CC -!- FUNCTION: ANCHORING PROTEIN THAT MEDIATES THE SUBCELLULAR

CC COMPARTMENTATION OF PROTEIN KINASE (PKA) AND PROTEIN KINASE C

CC (PKC).

CC -!- SUBUNIT: BINDS TO DIMERIC RII-ALPHA REGULATORY SUBUNIT OF PKA.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC. MAY BE PART OF THE CORTICAL

CC CYTOSKELETON.

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN ENDOTHELIAL CELLS, CULTURED

OY	338	---	QOEAEPTDQARISADVENKVELPELOVODLEASSSEKCAPLATIPEBKMAH-O	392
OY	354	QAHQOEPAESNAHEPRIASAETKEVLESEBQVSGOSPSEKAPLATEVFBKILEVHOE	413	
OY	393	VVAEVAHSTVEAKTEEBEEOGGGGEABGVVEGSESLPEPKLAEOVEVQOEAPAEELMKS	452	
Db	414	VVAEVAHSTVEBERNEQ-----KTEVEBPTAGSVPAEELVGMDAEQOEAPAEELVKL	465	
OY	453	REMCVSGDHTQULTDLSPEEKTLPKHPEGLVSEVEMLSQOEIKYQSGPLKLSSSGLK	512	
Db	466	KETVSGEDPTQGDADLSPEKVLSPKPEBVSEVEMLSQOEIMKYQSGPLKTLFTSTOLK	525	
OY	513	KLSSKKKKGKRGGGGGEPEPEYHITHTEPSADBDQKSSSSSPSEEBPTCLEKGLPE	572	
Db	526	KLSSKKKKGKR-GGGGESESEHTQVPADSPDSQOEKGBESSASSPEEBEITCLEKGLAE	584	
OY	573	APQGEAEBCGTSQGEKKRGKITPMASFKMWTPKKRVRPASEPKREBLEKVKATLSS	632	
Db	585	VQOQGEABEGATDQGEKKRGKITPMASFKMWTPKKRVRPASEDKEDBLKVYSATLSS	644	
OY	633	TDSTVSEMODEVKTVEBQKPEEPKRRVDTSVSWEALICVSSSKKRRKASSDDEGGPR	692	
Db	645	TESTASHEMOEEMKGSVEEPPEPKKRVDTSVSWEALICVSSSKKRRARRRSSDDEGGPK	704	
OY	693	TLGGDSHRAEBSKQKXAGCTDAPASTQGDQAQSSSPREPAGSSBEBGVTWMSFKRL	752	
Db	705	AMGGDHQKADBGADKETGTGILASQGHDPQSSSPQOAGSPTEBEGVSTWMSFKRL	764	
OY	753	VTPEKKSKSLKEEKAED---SSVEQLSTEIPRESRESVSIKKILPGRKKRAQKXQO	808	
Db	765	VTPEKKSKSLKEEBSBDSIANGSVHEISTPDTPEKESVSIKKILPGRKKRPPGKXQO	824	
OY	809	ATVEDSGPEINEBDDPNVPAVPLSEYNAREKME--AQNTBELPOLGAIVVSEELS	865	
Db	825	APVEDAGPTGANEDSDVPAPVPLSEYDAVEREKMEDAQOAGABQPEQKATEVSKELS	884	
OY	866	KTLVHTVSAVIGDTRATVSEERSSSMISASYTEPLEHTAGEAMPPEVETEKIILA-E	924	
Db	885	ESQVHMMAAVAAGOTRAATILIEERSPSMISASYTEPLEOVENBAALLTBEVLEREVIABE	944	
OY	925	ETPVLITOTLPEGKQADHMDWTSEVDTSEAVTATEPSALRTBEVTEAGAEETTDWVSA	984	
Db	945	EPTVTEPELPENRKARBDIVYSABELTPRAVTAETAGLSEBGETBSAAEETTEMVSA	1004	
OY	985	VSQULTSDPTTEATPTQVEVESGVLDTSEERQTOAILQAVADKYKESQVAPAT-----	1038	
Db	1005	VSQULTSDPTTEATPTQVEBEGVPDIEBERQRTQVILQAVAKYKESQJLPTGTGPEDV	1064	
OY	1039	QTVQRTGSKALEKVEVEEDSVLASSEKQVMPKGPVQEGABHLAOGSETGATPES	1097	
Db	1065	LQPVQRA-----EAERPEEOAASGLKKETDVLVQAOZKATPTPTQKVGVGCTPES	1118	
OY	1098	LE-VPEVTADVDH---VATQOV-----IKLOOL-MQOAVAPASSELTLDSEMTNGSTPLAD	1147	
Db	1119	FEKAPQVOTESIESELVTTQOAFTLGAVGSQOEVMEOALPBPUSVEPTDSETDGSTPAD	1178	
OY	1148	SDTADGQOODETTISQDSKATAAVROSQVTEEEAATAQKEBPSTPLNNVPAOEHEGEPG	1207	
Db	1179	FDACGTTQOKOIEYIHEHNEVNASCTQSGGTAEAVPAQKERPAP-PSSEVPQBEETKEQSK	1237	
OY	1208	RDVLEPTQOELTAALVPVLAKTEVQOEGEVDMLDEKVK-----EBQEVFNHSG-PNS	1255	
Db	1236	EKVTEVALKGEETEAACKDQDALEQSAKSPSPSEVEEMVAVYERETKEAPHNABE	1355	
OY	1315	G-EQETAPREHEGTGYKPVLTLLMPSEBKGALGSIQGS-PSLPRQDQKAGIEVOVQSD	1377	
Db	1336	KLEHETAVTVSEBVSQOLLOTVANVPIIDGAKVSSLESGPPCLQGOEBAVCTKIXQVQSSB	1415	
OY	1373	TTVQVTEAENAEKVLI--ETVVISSETGSPCEVGANHLIPAKSSATGGMWTLQHAEDTVPLG	1430	

Db 1416 ASFTLTAAREEKKVGTANILETGETLEPAGAHVLEESSEKNEDFAHPGEDAVPTG 1475
 QY 1431 PESQARSPIIVTAPPESTLHPDLQGISASQERSEEDKPDAGPDADKSTAIKVL 1490
 Db 1476 PDCQAKSTPVSATTKKGLSSDLEGEKTTSLKWSDEVDEQVACQEV--KVSVAIEDL- 1532
 QY 1491 KAPEP--ILELESKNKVLNLTQAVDQPART-ETAPETHAYDSQTVQACRLDSREP 1547
 Db 1533 --EPENGILSETKSKLVNIIQTAVDQVTEETATEMLTSELOTQAHVIKADSQDAG 1590
 QY 1548 RCWTK-----MKDAKMKHPVPQPREDLQVLTVLEA 1577
 Db 1591 QTEKESGEEPPQASQADETPITSKESESTAVGQA 1625

RESULT 2

MAPB MOUSE
 ID MAPB MOUSE STANDARD; PRT; 2464 AA.
 AC P14873;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Microtubule-associated protein 1B (MAP 1B) (MAP1.2) (MAP1(X))
 DE [Contains: MAP1 light chain LC1].
 GN MAP1B OR MTAP1B OR MTAP5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 RX MEDLINE=90094539; PubMed=2480963;
 RA Noble M., Lewis S.A., Cowan N.J.;
 RT "The microtubule binding domain of microtubule-associated protein
 MAP1B contains a repeated sequence motif unrelated to that of MAP2
 and tau.";
 RL J. Cell Biol. 109:3367-3376(1989).
 CC -!- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
 CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
 CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
 CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
 CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
 CC STABILIZING MICROTUBULES.
 CC -!- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
 CC WITH MAP1A AND MAP1B PROTEINS.
 CC -!- DOMAIN: Has a highly basic region with many copies of the sequence
 CC KKEE and KKEI/V, repeated but not at fixed intervals, which is
 CC responsible for the binding of MAP1B to microtubules.
 CC -!- PTM: LC1 IS COEXPRESSED WITH MAP1B. IT IS A POLYPEPTIDE GENERATED
 CC FROM MAP1B BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
 CC BOTH MAP1A AND MAP1B. IT INTERACTS WITH THE AMINO-TERMINAL REGION
 CC OF MAP1B.
 CC -!- SIMILARITY: TO MAP1A.
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X51396; CAA35761.1; --
 DR PIR; S07549; QRMSP1.
 DR MGI; MGI:1306778; Mtap1b.
 DR InterPro; IPR000102; MAP1B_neuraxin.
 DR Pfam; PF00414; MAP1B_neuraxin; 10.
 DR PROSITE; PS00230; MAP1B_NEURAXIN; 7.
 KW Microtubules; Repeat; Phosphorylation.
 FT CHAIN ? 2464 MAP1 LIGHT CHAIN LC1.
 FT REPEAT 1874 1890 MAP1B 1.

FT REPEAT 1891 1907 MAP1B 2.
 FT REPEAT 1908 1924 MAP1B 3.
 FT REPEAT 1925 1941 MAP1B 4.
 FT REPEAT 1942 1958 MAP1B 5.
 FT REPEAT 1959 1975 MAP1B 6.
 FT REPEAT 1993 2009 MAP1B 7.
 FT REPEAT 2010 2026 MAP1B 8.
 FT REPEAT 2027 2043 MAP1B 9.
 FT REPEAT 2044 2060 MAP1B 10.
 FT DOMAIN 589 787 KKEE AND KKEI/V REPEATS).
 SQ SEQUENCE 2464 AA; 270408 MW; FBD3DD99CFBDA87 CRC64;

Query Match 6.2%; Score 501; DB 1; Length 2464;

Best Local Similarity 21.1%; Pred. No. 2.4e-10;

Matches 387; Conservative 266; Mismatches 648; Indels 532; Gaps 86;

QY 45 ADPATKL-----PQKQGLSSVNGVAEQGDVH-----VQENQEQEVEVDEVDVQRE 93
 Db 478 ANPAEKIIRVLPFGNSTOYNILEGLEKHLKHLDFLKOPLATQKDLTGQVPTPPVKQVKLQ 537
 QY 94 SEDVREKDRVEEAMNSTAVEDITDKQGETSEI-----IEQIPASENNVEEMVQPAESQ 148
 Db 538 RADSRSE---LKPATKPVASKSVRKESKEETPEVTKTSQVEKTPKVESKE-----584
 QY 149 ANDVGFKVFVKFVGFKFTVKDKNEKSDTVQLLTVKKDEGEAGASVAGDHOEPPSVETA 208
 Db 585 -----KVLVKDKPKVTKSPSVT-----EKEVSKSEQSP-VKAE 619
 QY 209 VGESASKESELKQSTEKQEGTLKQSSSTEIPLOAESDQAAEEAKDEGEVEKQEKPTKS 268
 Db 620 VAEKQATESKPKVTDKK---VVVKE-----IKTLEEKKEEKPKKVEVKEDKTPFKK 669
 QY 269 PESP-SSPVNSSTTSFFKFFTHGWAGWRKTSFKKSKEDDLTAETAKRQEAQKVDDEE 327
 Db 670 DKPRKEEVKKIKKEIKK-----EERKELKVEVKKTPDKAKKVEKKEKK---EVK 720
 QY 328 KEKTEPASEEQPAEDTDOARLSADYKVELPLEDQVGDLEASSEKCAPLATEVFDE-- 385
 Db 721 KEEKPKKEIKKIDIKKSTPQSDTKPS-ALKPKVAKKEESTKKE--PLAAGKLKDKG 777
 QY 386 -----KWEAHQEVVAEVHVSTVEKTEEEEOGGGEAGGVVVEGTGESLPPEKLAEPQV 440
 Db 778 KVKVTKKEGKTTTAAATAVGTAAATTAAVVAAAGIAASGPFVKELEAER---SLMSSPEDLT 834
 QY 441 QBAEPAAELMKREMCVSGDHTQLTDLTSPEEK---TLPKHPEGIVSEVEMLSQRIKV 497
 Db 835 KQFER-----LKAEEIDVAKDIKQLELEDEEKLEKQETQGEAYVIQKETEVSQSAE--- 887
 QY 498 QGSPLKLFSSSGLKSLGKQKGRGGGDBEPGEYQHIHTESPESADEQKGESSASP 557
 Db 888 --SP-----DEGITTEGE-----GECEQTPEELEPVEKQGV--DIEKFEDEGAGF 930
 QY 558 EPEPETTCLKPLAPQDGEAEAGTSDGKKRKGITPWASFKQKQVTPKRVRRPSESD 617
 Db 931 EBSSET-----GDYBEKABTEAEAEPEDEGEDNASG-----SASK-----HSPTEDD 972
 QY 618 KEEELE-----KVKSATLSSTDTVSEMDVETKTVGEEKPEEPKRRVDTSVSWEALICV 672
 Db 973 ESAKAEADVHLKEKRESVVSGDDRAEEDMDVLEKGEAFQSEEGEED-----1021
 QY 673 GSKKRRKARAKSSDDEG-GRPTLGGDSHRAEASKDKAGTDAVPASTQEQDQAQSSS-- 729
 Db 1022 -----KADEAREGEYEPDKTEAEDYVMAVADKAAEAGV-----TEEQYGLGTSK 1067
 QY 730 -----SPEPAGS-----PSEGEVSTWESPK-----RLVTPRKSKSKLEKAEAD 769
 Db 1068 QPGIOSPSREPASIHDETLPGGSESEATADENREDQPEEFTATSGVYQTSTIETSSRP 1127
 QY 770 SSVQELST-----EIPSRSEESVSIKPF-----IPGR 798
 Db 1128 TPMDEMSTPDVMSDETNNETESPSQBFVNITKYESSLYSQEYKPAVASPGLSEGSK 1187

```

QY 799 KKRADCKQOATVEBSPVEINEDDNVAVPLSEYNNAVERKMEAQONTLPOLLGAV 858
DB 1188 TDATODCKDNASASTISPPSSMEEDFKSKA--LRDAYCSEEEKELASALDIKD----- 1240
QY 859 YVSEELSKTLVHTVAVIDGTRAVTSVEBSPMSIASVTEPLEHT-AGE-----AMPP 912
DB 1241 -VSDE-----RLSPA-----KSPS-LSPPSPPIEKTPLGERVSNLSRP 1278
QY 913 VE-EVTEKDIIEETPVLTQTLLEGKADHDWVTSEVDTSEAVTAT-----E 959
DB 1279 NEIKVAGEBARSVSPGVTAQVAVEEHCAEPEKTLLEVSPSQVTSAGHTPYQSPDDE 1338
QY 960 TSEALRTE-----EVEEASGAEETT-----979
DB 1339 KSSHLPTEVSENAQAVPSFEFEADENERASLSPMDEBPDSPEPVKSLPLSPPL 1398
QY 980 -----DMVASVSQL-----TSP-----DTTEATPVQEVES-GVLDTEBEROT 1018
DB 1399 LGSSEPYEDFLSADSKVLRGRSSEPECKNGKQGFDPDRSPVDLTSTGLYDKQEKST 1458
QY 1019 QALIQAVADKVEE--SQVPATQTVORTGSKALKKEVEEEDSEVLASEKEKDVMPKGPVQ 1077
DB 1459 GFL-----PIKEDFGEPEKKTSDVETMSSQSALALDERKLGIDV--SPQIDVSGSFK 1510
QY 1078 EAGAEHLAOGSETGQ-ATPESLVEPEVT-ADVDHATCOVIKLOQMEQAVAPESSETLT 1135
DB 1511 EDTKMSISEGTVSDKATPDECAVDYTHMGVAVS-----STAVALTSSFP 1559
QY 1136 DSFTNSTPLADSDTADGQODETISQDSKATAAARQSOVTEBEAATAQKE----- 1188
DB 1560 EPTTDVDSVSLHAEV--GSHSTFVD--DLSLSVSVQTPPTFOTEMSPKEKCPRMSI 1615
QY 1189 -----PSTLPNNVPAQOEHEBGRDV-----LEPTQOELTAAAVPY 1225
DB 1616 SPDFSPKTAKSRTPVQDHRSEQSSMSIERGQSPHESFAMDRSPDPHTLGAS---V 1672
QY 1226 LATVEGQSEVWMLDGEKKYQEVFVHSGPMSOKAADVTYP---SEWGVAGCQEKES 1282
DB 1673 LHITENGPT-EVDY---SPCDIDSSLSHKIPTEEPS-YTQNDLSELSVSGVAPSPS 1727
QY 1283 T-----EVOSLSLEEGEMETDVEK-EKRE 1305
DB 1728 TSSAHTPSQIASPLQEDTLSDVPPRPMSLYASLASEKVSLEGKLSPKSDISPLTPRE 1787
QY 1306 TKP-----EQVSEEGEETAPEHEGTYGKV-----LTLTD 1336
DB 1788 SSLVSPGFSDSIAAKETA-AHQASSSPPIDAATAPYGFRRSMLFTMQHLLNND 1846
QY 1337 MPSS-----ERKALG-----SLGSPSLPDDAKGCTIEVYQSLDT--TYTQTA 1379
DB 1847 LTTSSVEKDSGGKTPGPFNYAYOKPEVAAGSPDEEDYD---YESQEKTRTDVVRVYV 1902
QY 1380 EAVEKXILETVVISIETGESPPCVGAHLLPAKKSAT--CGHWTLOHAEDEVPLAPESQAS 1437
DB 1903 EKTERTIKSP--CDSGSYSTET-----EKTIKPEBGGYTEIETIKTTRT-PEESGYS 1952
QY 1438 IPIVTPAPESTLAPDLOGEISASQERSEEE-----EDKPDAGPD-ADGKESTALE 1487
DB 1953 YEL-----SEKTTTPREVSG-YTEKTERSRRLDDISNGVDJDEGHTLGDGYSYETT 2007
QY 1488 KVLAEPEILLESKSNKIVLANVIOAVADQAPARTETAPEHTAY---DSQTOVPACRLD 1542
DB 2008 EKITSFP--ESESYS-----YETSTKTSRSPDTSAYCAYETEMKITRTPQASTY 2053
QY 1543 SRE-PNRCWTMKMDAKMKHPVQPREDLQVLT 1574
DB 2054 SYETSRCYTTERK--KSP-SEARQDVLDCLV 2081

```

```

DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-OUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
RL J. Cell Biol. 114:241-253 (1991).
RN [2]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeli E., Bennett V.;
RT "440-kD ankyrinB: structure of the major developmentally regulated
RL domain and selective localization in unmyelinated axons.";
RN [4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
RA Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
RL ankyrin gene.";
RN [5]
RP Genomics 10:858-866 (1991).
CC -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC elements. Also bind to cytoskeletal proteins.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS, 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CC CELLS THROUGHOUT THE BRAIN.
CC -!- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X56957; CAA40278.1; -.
DR EMBL, X56958; CAA40279.2; -.
DR EMBL, Z26634; CAA42644.1; -.
DR EMBL, M37123; AAA62828.1; -.
DR PIR, S14533; S14533.
DR PIR, A39643; A39643.
DR PIR, B39643; B39643.
DR PIR, S14569; S14569.
DR HSSP, P42771; 1DC2.
DR Genew; HGNC:493; ANK2.
DR MIM; 106410; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.

```



```

QY 930 TQTLBEGKADHDMVTSEVDFTSEAVTATETSEALTE-----EYTEAS 973
Db 2685 -YTFKNNEDTQEEBPGKSEBKESESHLA-EDRHAVSTEADNRSDYKLNDDTQPKICDGH 2742
QY 974 GAETITDWSA--VSQLTSP-----DYTEATPVQEVESG 1007
Db 2743 GCEANSPSSARPVSSGLSPGDDVDQPVLYKESLALQGTHEKOTBEELDVSAEESP 2802
QY 1008 VLDTEEBERQOTAIQ--AVADKVEESOVPAQTQVORFG-SKALEKVEEVED----- 1058
Db 2803 QADCESESSESSSSLPCHVSEBKELEDIEDISATSSIQKTEVTKTDTEFNLPDPCPSQDS 2862
QY 1059 -----SEVLASEKEKVMKGPVQEAHMLAQGSQTGATP-----E 1096
Db 2863 STTTQDTRFSDMVPVSDLLENDDIYDPQTSPYENVPSSQFSSSESKCTQDANTTSH 2922
QY 1097 SLEBEV--TADVDHATC-----QVILQOLMEQAVAPESSELTIDSE 1138
Db 2923 SSEVSVTITSPVEDVVAVSSSGTVLSKESNPEGODIMEQSLESTL---WEMQSDSV 2978
QY 1139 TNGSTP-LADSDPTADGTQDETIDSQ--DSKATAVR-----QSQVTEE---AATA 1184
Db 2979 SSSFEPTWATTTVGEQISKVITKTVDSDSWSSEIRDEDAFEARKEEKGKIFGLNV 3038
QY 1185 QKEBPTLPNNVPAGEHEBEPGRDVLPTQEL-TAAVPIVAKT---EVQGBEVDW 1239
Db 3039 DQOSQGTDTTTPAR-----TPTBETPTSEONPFLFQGKLFEMTRSGAID- 3085
QY 1240 LGEKKEKEQEVFVHSGPSOKADVTYDSEVW-GVAGCOEKE-STEVOSLSLEGEMET 1297
Db 3086 MKRKSTADSFHFPOIG---QESRETLSEDKVKGATGADPLPLETSASLSLSEKETV 3142
QY 1298 DVEKEKRETKPEQVSEBEGQ-----ETAAPEHGTGYGKPVLTLDMPSS 1341
Db 3143 D--DEADLLPDSVSEBEVEIPASDAQLSQMGISASTETPKEAV---SVGKDLPTVQ 3196
QY 1342 RKGALGSLG--SPSLPDQKACIEVQVQSLDTTQTAHAEKVIETVIVISETG-ESP 1398
Db 3197 TQD-TPPLGVAQKQISCPDSEPA--VQVQLDPTLTRS-----VYSDRDSDSP 3241
QY 1399 ECVGAHLPAKSSATGCHWTLOHADTVPLGP-----ESQAESPIPI---ITVPAPE 1447
Db 3242 DS-----SEEBQKSVI-----EIPYAMENVPTRESK-K-IPRTMPTSTPAP 3284
QY 1448 STLHPDLOGEISASQERSEEBDKPDAGPDAGKESSTALEKVL-KAPELLEISKSKI 1506
Db 3285 SA---EYESSVSEDFLSVDEEKAD--EAKPKSKLPYKVPLORVQEQSLDLDTSVQKT 3338
QY 1507 VL---NVQTAVDQFARTETAPETAHAYSQVAPACRLDSREPNCWTMKMDAKKHRY 1562
Db 3339 VAPQGDMASTIAPDNRSKES--DASLSDKTKCPV-----KTRSYETETESRER-- 3387
QY 1563 PQPREDLQVLTLEAWQPRKCLPRLQKA 1592
Db 3388 ---AEELF-LESEBEGATRPKILTSRLPVKS 3413

```

```

RESULT 4
MAPB_HUMAN
ID MAPB_HUMAN STANDARD; PRT; 2468 AA.
AC P46821;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 1B (MAP 1B) [Contains: MAP1 light chain
LC1].
GN MAP1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC Tissue=Fetal brain;
RX MEDLINE=95104835; PubMed=7806212;
RA Lien L.V., Feener C., Fischbach N., Kunkel L.M.;
RT "Cloning of human microtubule-associated protein 1B and the
RL identification of a related gene on chromosome 15."
CC Genomics 22:273-280(1994).
CC
CC -!- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
CC STABILIZING MICROTUBULES.
CC -!- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS.
CC -!- DOMAIN: Has a highly basic region with many copies of the sequence
CC KKEE and KKEI/V, repeated but not at fixed intervals, which is
CC responsible for the binding of MAP1B to microtubules.
CC -!- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated
CC from MAP1B by proteolytic processing. It is free to associate with
CC both MAP1A and MAP1B. It interacts with the amino-terminal region
CC of MAP1B (by similarity).
CC -!- SIMILARITY: TO MAP1A.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L06237; AAA18904.1; -.
CC DR Genew; HGNC:6836; MAP1B.
CC DR MIM; 157129; -.
CC DR InterPro; IPR000102; MAP1B neuraxin.
CC DR Pfam; PF00414; MAP1B neuraxin; 10.
CC DR PROSITE; PS00230; MAP1B_NEURAXIN; 6.
CC KM Microtubules; Repeat; Phosphorylation.
CC FT CHAIN ? 2468
CC FT REPEAT 1878 1894 MAP1B 1.
CC FT REPEAT 1895 1911 MAP1B 2.
CC FT REPEAT 1912 1928 MAP1B 3.
CC FT REPEAT 1929 1945 MAP1B 4.
CC FT REPEAT 1946 1962 MAP1B 5.
CC FT REPEAT 1963 1979 MAP1B 6.
CC FT REPEAT 1997 2013 MAP1B 7.
CC FT REPEAT 2014 2030 MAP1B 8.
CC FT REPEAT 2031 2047 MAP1B 9.
CC FT REPEAT 2048 2064 MAP1B 10.
CC FT DOMAIN 589 790 LYS-RICH (HIGHLY BASIC, CONTAINS MANY
CC KKEE AND KKEI/V REPEATS).
CC SQ SEQUENCE 2468 AA; 270618 MW; 540839CDBD09461 CRC64;

```

```

Query Match 6.1%; Score 492; DB 1; Length 2468;
Best Local Similarity 20.3%; Pred. No. 4.9e-10;
Matches 368; Conservative 266; Mismatches 694; Indels 482; Gaps 75;

```

```

QY 45 ADPATKL-----PQKNGQLSSVNGVAEQDVH-----VQENQSGQEEVVDVQVQRE 93
Db 478 ANPAKRIIRVLEFGNSGYNILEGELKXLDPLKQPLATQKDLTGCVPTPVVKQTKLKQ 537
QY 94 SEDVEKQVVEEMANSTAVEDITQGGQETSEITQIPASENNVEVWQPAESQANDVG 153
Db 538 RADSRSS--LKPAPKPLPSKSVKRESKEFREV-----TKVNHVE--KPKVSKSKE-- 584
QY 154 FKVKFVGFKFTYKQDKNKSQDLYQLLTVKQDEBGBAASVAGADHDQPS-VEFVAGES 212
Db 585 -----KVMVKDDPVKKTETKPSVTEKE-----VPSKPEPSPVKAELAE- 622
QY 213 ASKESELKQSTEGKGLKQEOSSTEIRPLQAESDQAEBAKDEBEERQKQPTSPSPSP 272
Db 623 -----KQATDVKPKAKAKETVKKETVYKPEBDKKEEKKPKGKVAKKEDKTPIKKEKP 675

```


RT	"Neuraxin: a novel putative structural protein of the rat central nervous system that is immunologically related to microtubule-associated protein 5."
RT	EMBO J. 8: 2879-2888(1989).
RL	[4]
RN	DEVELOPMENTAL STAGE, AND PHOSPHORYLATION.
RP	MEDLINE=97405699; PubMed=9260743;
RX	Ma D., Nothias F., Boyne L.J., Fischer I.;
RA	"Differential regulation of microtubule-associated protein 1B (MAP1B) in rat CNS and PNS during development.";
RT	J. Neurosci. Res. 49:319-332(1997).
RL	- FUNCTION: The function of brain MAPs is essentially unknown. Phosphorylated MAP1B may play a role in the cytoskeletal changes that accompany neurite extension. Possibly MAP1B binds to at least two tubulin subunits in the polymer, and this bridging of subunits might be involved in nucleating microtubule polymerization and in stabilizing microtubules.
CC	- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate with MAP1A and MAP1B proteins.
CC	- TISSUE SPECIFICITY: Nervous system (spinal cord, brain stem, cerebellum and cerebrum). Not expressed in liver, spleen, kidney, heart or muscle.
CC	- DEVELOPMENTAL STAGE: In cerebral cortex, spinal cord and sciatic nerve levels are high early in development but decrease during postnatal development and are low in adults. In dorsal root ganglia levels remain high throughout development.
CC	- INDUCTION: By nerve growth factor.
CC	- DOMAIN: Has a highly basic region with many copies of the sequence KKKE and KEK1/V, repeated but not at fixed intervals, which is responsible for the binding of MAP1B to microtubules.
CC	- PFM: LC1 is coexpressed with MAP1B. It is a polypeptide generated from MAP1B by proteolytic processing. It is free to associate with both MAP1A and MAP1B. It interacts with the amino-terminal region of MAP1B (By similarity).
CC	- PTM: Phosphorylated.
CC	- SIMILARITY: TO MAP1A.
CC	- CAUTION: A C-terminal fragment of this protein (residues 1597 to 2459) was originally described as neuraxin in Ref.3.
CC	-----
CC	This SMS-S-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC	-----
DR	EMBL; U52950; AAB17068.1; -
DR	EMBL; X60370; CAC16162.1; -
DR	EMBL; X16623; CAA34620.1; ALT. SEQ.
DR	PIR; S06017; S06017.
DR	InterPro: IPR000102; MAP1B neuraxin.
DR	pfam: PF00414; MAP1B neuraxin; 10.
DR	PROSITE; PS00230; MAP1B_NEURAXIN; 8.
KW	Microtubules; Repeat; Phosphorylation.
FT	CHAIN ? 2459
FT	REPEAT 1869 1885
FT	REPEAT 1886 1902
FT	REPEAT 1903 1919
FT	REPEAT 1920 1936
FT	REPEAT 1937 1953
FT	REPEAT 1954 1970
FT	REPEAT 1988 2004
FT	REPEAT 2005 2021
FT	REPEAT 2022 2038
FT	REPEAT 2039 2055
FT	REPEAT 2056 2072
FT	REPEAT 2073 2089
FT	REPEAT 2090 2106
FT	REPEAT 2107 2123
FT	REPEAT 2124 2140
FT	REPEAT 2141 2157
FT	REPEAT 2158 2174
FT	REPEAT 2175 2191
FT	REPEAT 2192 2208
FT	REPEAT 2209 2225
FT	REPEAT 2226 2242
FT	REPEAT 2243 2259
FT	REPEAT 2260 2276
FT	REPEAT 2277 2293
FT	REPEAT 2294 2310
FT	REPEAT 2311 2327
FT	REPEAT 2328 2344
FT	REPEAT 2345 2361
FT	REPEAT 2362 2378
FT	REPEAT 2379 2395
FT	REPEAT 2396 2412
FT	REPEAT 2413 2429
FT	REPEAT 2430 2446
FT	REPEAT 2447 2463
FT	REPEAT 2464 2480
FT	REPEAT 2481 2497
FT	REPEAT 2498 2514
FT	REPEAT 2515 2531
FT	REPEAT 2532 2548
FT	REPEAT 2549 2565
FT	REPEAT 2566 2582
FT	REPEAT 2583 2599
FT	REPEAT 2600 2616
FT	REPEAT 2617 2633
FT	REPEAT 2634 2650
FT	REPEAT 2651 2667
FT	REPEAT 2668 2684
FT	REPEAT 2685 2701
FT	REPEAT 2702 2718
FT	REPEAT 2719 2735
FT	REPEAT 2736 2752
FT	REPEAT 2753 2769
FT	REPEAT 2770 2786
FT	REPEAT 2787 2803
FT	REPEAT 2804 2820
FT	REPEAT 2821 2837
FT	REPEAT 2838 2854
FT	REPEAT 2855 2871
FT	REPEAT 2872 2888
FT	REPEAT 2889 2905
FT	REPEAT 2906 2922
FT	REPEAT 2923 2939
FT	REPEAT 2940 2956
FT	REPEAT 2957 2973
FT	REPEAT 2974 2990
FT	REPEAT 2991 3007
FT	REPEAT 3008 3024
FT	REPEAT 3025 3041
FT	REPEAT 3042 3058
FT	REPEAT 3059 3075
FT	REPEAT 3076 3092
FT	REPEAT 3093 3109
FT	REPEAT 3110 3126
FT	REPEAT 3127 3143
FT	REPEAT 3144 3160
FT	REPEAT 3161 3177
FT	REPEAT 3178 3194
FT	REPEAT 3195 3211
FT	REPEAT 3212 3228
FT	REPEAT 3229 3245
FT	REPEAT 3246 3262
FT	REPEAT 3263 3279
FT	REPEAT 3280 3296
FT	REPEAT 3297 3313
FT	REPEAT 3314 3330
FT	REPEAT 3331 3347
FT	REPEAT 3348 3364
FT	REPEAT 3365 3381
FT	REPEAT 3382 3398
FT	REPEAT 3399 3415
FT	REPEAT 3416 3432
FT	REPEAT 3433 3449
FT	REPEAT 3450 3466
FT	REPEAT 3467 3483
FT	REPEAT 3484 3500
FT	REPEAT 3501 3517

```

SQ SEQUENCE      2459 AA; 269497 MW; 2E3F6872DEDB8BA2 CRC64;
Query Match          5.8%; Score 470; DB 1; Length 2459;
Best Local Similarity 20.2%; Pred. No. 2.8e-09;
Matches 328; Conservative 234; Mismatches 602; Indels 458; Gaps
62;

QY 45 ADPAKLU-----PQNGGLSSVNCVAAGGDH-----VGEENQEGEEVVDDVGORE 93
Db 477 ANPAEKITRLVPFGNSTIYNILBGGJEKKLHIDFLKOPLATOKDITGVSTPPVKOVKLKQ 536
QY 94 SEDVE--KDRVEEMAANSTAVEDITDKGQEETSEIIIEQIPASENNVEWQPABSOAND 151
Db 537 RADSESJLKPATKPLSSKS-----VRKSKREAP-----ATASQYEKTPKYESKE-- 583
QY 152 VGFKAIVKFVFGRKFTVKDKNKEKDVTVOLLTIVKKDBEGGAASVAGDHQEPSETAVGE 211
Db 584 -----KVIVKKDKDGKVESKPSVT-----EKEVPSKEEQSP-VKAEVAAE 621
QY 212 SAKSSELKQSTKEQEGTLKQEOSTELPILOAESDOAAEEAOKGBEKOKEPTKSPES 271
Db 622 KAATESKRKYTKD--VKKE-----IKTPBEKKEEPKPEVAKGEDKTPKDKOEK 671
QY 272 PSSPNSETTSSFKKFTFHGAWRKRTKSPFKSKEDDETLEAKREKQEAKEYDEBK--E 329
Db 672 PK--KEBAKKEIK-----EIKKEKEKEL-KEVKKEPTLKNAKKEVKDE 714
QY 330 KTEPASSEOEPADTDQARLSADYKVELPLEDOVDIGLEASSSEKCAPLATEVPEKMEA 389
Db 715 KKVEVKEKEPEKKEI--KKISKDIK-STPLSD-----TKKPALKRKPVAKEBPT 762
QY 390 HOEVAAEHVSVTEKTEEBEGGGGGAEGGVVEGTGESLPPEKLAEPQEVQEAABEL 449
Db 763 KEPIPAAGLKDKGVAVKIYKKGKTEETAATAVGAIAAAAAAGA-----ASPAPAKEL 815
QY 450 MKSRMCVSGGDHTQ-LTDLISPEEKTLPRKHPEGIASEVMLSOERIYKQSPPLKLFSS 508
Db 816 EAKERLMSSPELUTCDPBEHLKAEIEDVAK--DIKPOLLELBDEBKUL----- 860
QY 509 SGCLKLISGKKOKGKRGGGGDEBPGEYOIHITESPESADEQKGESSASSPEEPETTLEK 568
Db 861 -----ETEPGAAYIQKE-----TEVSK 878
QY 569 GELEAPQOGAEBGTTISOGEKKREGITTMASFKKMTFKKAVRPRESBDKEBLEKYSYA 628
Db 879 GSABSPDREGI---TTTGEGECE-----QTPr-----EELEPV--- 907
QY 629 TLJSSTDTSVSEMQDEVKTVGEEKQKEEPKRRVDTSVSMELICVSSSKRKARKASSDDE 688
Db 908 -----EKQYVDILEKPEDBAGFBESSAGDYIEKA----- 938
QY 689 GGPRTLGGDSHRABEASKDKREAGTDAVASTOBODQAQSSSPBPAGSPBSGEGVSWES 748
Db 939 -----ETFEABEPBEDG-----EDNVGSASKH--SPTDEBITAKABA 974
QY 749 FKRLVTPRKKSXSKLEKRAEDSVQOLST-ELEPBRESWWSIKKFTIGRRKKRADGOE 807
Db 975 DVH-IKERESYASADDRAEDMDALKEKGAEOSEEBG-----EEBEDRAE 1020
QY 808 QATVEDSGVEINEDDPNPVAVPVLSEFNVAVERKMEAQGNTELPOLLGAVVVSBELSKT 867
Db 1021 DAREEDHEPDKTEAAD-YMAAVVDKRAERGCTEDQYDFLG---PAKPGVOVSPREPAS 1076
QY 868 LVHTVASVAVIDTRAVTVEER---SPSNISAS-VTEPLEHTAGEAMPVVEVTEKDI 922
Db 1077 SJHDLETPEGSSSEANTASDENREDQPEEFATTSYTOSTIEISEPPIPMDEMSTRPDM 1136
QY 923 ABEET-----PYLT-QTLPECK-DAMDMMTSEVD 949
Db 1137 TDEINNEETESPQEFVNITKYESSLYSQEYSKPVVAFSGNLSDSKXTDATDGR----D 1191
QY 950 FTSBNAVTAETS-----EALTREVTSASGAEEFTDWVS-----AVSQLT 989

```

```

Db 1192 YNASASTISPPSMEEDKFSKALRDAYRPEETDVKTAGELDIKDVSDERLSPAKSPSL 1251
QY 990 DSPDTTEATPVORVSGVLTDEERQTOAILQAVADKVKESQVPATQVORTGSKAL 1049
Db 1252 PSPSPIEKTPLGERSVNFSLTNE-----IKASAEGETATVWSPGVTOAV 1297
QY 1050 -----EKVEEVEDSEVLASEKEKVMKPGVQVQEAHAQSGSEGTQATPESLEV 1100
Db 1298 VEHCASPEEKTUWVSPQSOTVSGAGHTPYQSPDEKSSHLPTEVTENAAQVPSVFEF 1357
QY 1101 PEVTADVDHVCATC-----QVTKLQQLMEQAVAPESSETLTDTSETNGSPFLADSDTADCTQ 1156
Db 1358 TEAKDENERSSISPMDEPVDSPSPKVLPLSPPLSGESAYEDFLSADKALGRS 1417
QY 1157 DETIDQDSKATNAVRQSVQTEBEAATAQKEPSTLPNNVPAQEEHGEPRGRVLEPTQ 1216
Db 1418 ESPFEKNGKQGFSDKESPVSDLTSLDYQDKBEKAGFTPIKEDFSPEKKASDAEIMSS 1477
QY 1217 ELTAAAVPVLAKEVGOEG---EVDMLDGEKVKEEQEVFVHSGPNSOKA-----ADVT 1266
Db 1478 QSALA-----LDERKJGDGSPTOVDVSGFSGFKEDTKMSISEGTVSDKATPVDEGAEDT 1533
QY 1267 YDSEVMVAGACQ-----EKESTEVEQSLSLEGE----- 1294
Db 1534 Y-SHMEGVASVSTASVATSFPEPTDDVSPSLHAEVSGSPHSTEVDLSVSVVQPTTF 1592
QY 1295 METDVKREKE-----TKPEQVSEEGEQTAAPEHEGYGKPVLTLDMPSSERKALG 1347
Db 1593 QETEMSPSKECPKPMGSPISPPDFSPKTSKRTPVQDHRSPQSSMSIFGQESPEHSLAM 1652
QY 1348 SLGGSPLPDQKAGCT-----EVQVQSLDTTQVTOAEAV-----EKVIET 1388
Db 1653 FSRQSPDHPTVG-AGMLHITENGTEVDYSPSDIQDSSLHKIPPTPEPSYTDNDLSEL 1711
QY 1389 VITSETGESPECVCAHLLPAEKSSATCGHWTLOHAEDTVPGLPESQAESPIIIVTPAPES 1448
Db 1712 ISVSQVEASPTSSAH-TPSQIAS-----PLQEDTLSD-----VVPDRM 1750
QY 1449 TLHPDLOGE-ISASORER-----SEEDKPDAGP-----DADGKESTAIEKVLKA 1492
Db 1751 SLVASLASEKQVSLGEGKLSPKSDISLTPRESSPTVSPGFSSTSGAKESTAAYQTSSS 1810
QY 1493 EP 1494
Db 1811 PP 1812

RESULT 6
MAPA_RAT
AC P34926;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 1A (MAP 1A) [Contains: MAP1 light chain
  LC2].
GN MAPA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92355629; PubMed=1379599;
RA Langkopf A., Hammarback J.A., Mueller R., Vallee R.B., Garner C.C.;
RT "Microtubule-associated proteins 1A and LC2. Two proteins encoded in
  one messenger RNA."
RL J. Biol. Chem. 267:16561-16566(1992).
CC -!- FUNCTION: Structural protein involved in the filamentous cross-
  bridging between microtubules and other skeletal elements.
CC -!- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate
  with MAP1A and MAP1B proteins.

```

```

CC -!- TISSUE SPECIFICITY: BRAIN, HEART AND MUSCLE.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED LATE DURING NEURONAL DEVELOPMENT
  APPEARING WHEN AXONS AND DENDRITES BEGIN TO SOLIDIFY AND STABILIZE
  THEIR MORPHOLOGY.
CC -!- DOMAIN: THE BASIC REGION CONTAINING THE REPEATS MAY BE RESPONSIBLE
  FOR THE BINDING OF MAP1A TO MICROTUBULES.
CC -!- PTM: VARIOUS SERINE RESIDUES MAY BE PHOSPHORYLATED BY CAMP KINASE.
CC -!- PTM: LC2 IS COEXPRESSED WITH MAP1A. IT IS A POLYPEPTIDE GENERATED
  FROM MAP1A BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
  BOTH MAP1A AND MAP1B.
CC -!- SIMILARITY: TO MAP1B.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M83196; AAB48069.1; -.
CC PIR; A43359; A43359.
CC CHAIN 22465 2774 MAP1 LIGHT CHAIN LC2.
CC MICROTUBULES; Repeat; Phosphorylation.
CC FT DOMAIN 309 496 LYS-RICH (BASIC).
CC FT DOMAIN 336 541 11 X 3 AA REPEATS OF K-K [DE].
CC FT REPEAT 336 338 1.
CC FT REPEAT 415 417 2.
CC FT REPEAT 420 422 3.
CC FT REPEAT 424 426 4.
CC FT REPEAT 427 429 5.
CC FT REPEAT 431 433 6.
CC FT REPEAT 436 438 7.
CC FT REPEAT 440 442 8.
CC FT REPEAT 444 446 9.
CC FT REPEAT 449 451 10.
CC FT REPEAT 539 541 11.
CC SQ SEQUENCE 2774 AA; 299526 MW; 3DEF74427BA9D7D7 CRC64;

Query Match 5.4%; Score 433.5; DB 1; Length 2774;
Best Local Similarity 22.0%; Pred. No. 5,5e-08;
Matches 388; Conservative 231; Mismatches 625; Indels 523; Gaps 89;

QY 103 VEEMAANSTAVEDIT-KDQGE-----ETSEIEIQIPASENNVMVQ---PAESQA 149
Db 216 MQWAGNSKAKTGIVLANGKEAEISVPYLTSITALLVWLPA--NPTEKIVRVLPFGNAPQ 273
QY 150 NDV--GFKV--FKFVGFKFTVKDKVKSQSDTQVLLTVKKDEGEAEASVAGDHQEPSV 205
Db 274 NKILEGLEKRLHLDFLRYPVATQKDLAAGVAPANLKPSPKIKHRADSKESLKA-----AP 327
QY 206 ETAVGESASKESELKQSTKQSGTLKQEOSSTEIPLQAESDQAABEAEKDEGEKEKEP 265
Db 328 KTAVSKLAKREVLLEGAKESLAKELAKT-----EKKAKEP 366
QY 266 T-KSPSPSPSP-----VNSSETTSPFKFTTHGWAGWRKKTSPFKSKEDDLETAEKKEQE 319
Db 367 SEKPPKPSKSERVGSSEALKAERLIKDKAG-----KXHLKEKISKLEKKDKE 419
QY 320 AKKVDDEEKE-KTEPASERQEPADTDQARLSADYKVELPLEDQVGDLEASSEEKCAPL 378
Db 420 KKEIKKERKELKEGRKEKKDAKKDKR-----KDTKPEVKLSKPLDKPF 467
QY 379 ATEVFDE-----KMEAHQEVVAEVHVSIVTEKEEEOGGGGAEGVVVGTSGL 428
Db 468 TPEVRKTLVYKAKAPGRKVDKGRARAARGEKELSESPRTPPAQKG---AAPPAAVSGHREL- 523
QY 429 PPEKLAEPQEVQEAEPABELMKSRMCMVSGGDHTQTLTDLSPEEKTLP----- 476
Db 524 ---ALSSPDLTQD---FEELKREER-----GLLAQRDTGLGEKPLPADATEQGHPSAA 572
QY 477 ---KHPEGTIVSEVMLSSQERIKVQGSPLKLFSSSGLKKLSGKKQKRGKGGGDEEPE 533

```

```

Db 573 IQVTPSGVLBGEHVEREKEY-VPDSPDK-----GSTNRGPDGSAEYEXEKE 620
Qy 534 YOHITHTEPSADEQKSSASSPEPE---ETTCLKEKPLEA-----POGEAEGETTS 585
Db 621 -----TWEERKQREALGPENTTAARESEAEVEDVIEKALEEMETHPSDEGEETAE 676
Qy 586 DGEKK--REGITPMASFCKMWTPKRVR-----RSESEKEELEKVSATLSSTDS 635
Db 677 SFYQKHTREAL-----KASPKREALGGRDLGFOGKAPKETASFLSLTATGATE 728
Qy 636 TVSEKODEV-----KTVGEOKPEEPKRRVDTSVSEWALICVSSSKRARRKSSSD 687
Db 729 HVSITQDETIPEYSETEQIISDEEIHDEPDER-----PAPRPFSTYTD 772
Qy 688 EGGPRTLLGDSHRAEASKDEAGTDAVPAS---TQEDDQAGSSSPPEAGSP-SEGEV 743
Db 773 LSGPEGPG-----PREAS---QAADSAVPASSSKTYGAPETELTYPPMVAAPLAEETHV 824
Qy 744 STWESFKRLVTPRKSKSKLEKAEADSSVEQLSTELPERESWVS-----IKKFIPIGR 798
Db 825 SSATS-----ITECDKLSSPATSVAEDOSVASLTA---PQTEBTGKSSLLDVTGSIPISSR 877
Qy 799 KKBAGKQEQALVEDSGPEINEDDPNPAVVPPLSEVYNAVEREKKEAOGNTELPOLLGA- 857
Db 878 TEATQGLDVPESAGTISPTSSLEBDKFKS--PCBEDPVSITGES-EKGETVGRGLSGEK 934
Qy 858 -----VYVSEELSKTLV-----HTVSAVAIDGTRAVTSVEER--SPSMISAVTE 900
Db 935 AVGKEEKYVNVSEKLSGQVAAVFGAPGHTLP--PGEPALGEVERCLSPDSTVYKMS 990
Qy 901 P-----LEHTAGAMPVEEVTEK----- 919
Db 991 PPPSGPPSAHATPHQSPVEDKSEPRDFQEDSGWGTGKSHPGVSKEDSEQTVKPGEEGT 1050
Qy 920 -----DIAETPVLVLTLP-----GDAHDWMTSEV 948
Db 1051 SEEGKPPTRSQADMPVSLGGQIGCTIQLLPEDKALIVETGAGNLAGTLPGEV 1110
Qy 949 DFTSEAVTATETSEALRTBEVTEASCAETTMVSAVSQUTDSPDTT--EATPVOEVESG 1007
Db 1111 RSTTEATPQODEVLR---FTDQSLSPEDASLSVLSV--SPDTTKQEAHP-----RS 1160
Qy 1008 VLDTEBEE-----KOTQALIQAVADKVKESQVPAQTVYVRTSKYL--EKVEE 1054
Db 1161 PCSLKEQOQHKMLMPVSPEDTQSL-----SFSEBS--PSKRTSLDISSKQLSPESLGT 1212
Qy 1055 VEDESEVLASEKQVMKPGVQGEAG--HLAQGS--ETGAT--PESLEVP----- 1101
Db 1213 IQFGELINLGSKEB---RGVPMKAEDDSCHLAPVISIPPHRATVSPSTDETPAGTLPBG 1266
Qy 1102 -----EYVADVDHATCQVILQOLMEQAVAPESSETLTDSETNGSTPL 1145
Db 1267 SFSHSALSVDRKSHPEITGPGHEMT-----SDSSLTKPSPLSSPA 1309
Qy 1146 AD-----SDTADGTOQDEITDSQSKATAVAVROSQVTEBEATLACK-----EPSTLP 1193
Db 1310 MEDLAVEMEGKAPGKEKEBELKSETRQ-----QKQILPEKVAVVDOLLIHOKGALDE 1364
Qy 1194 NNVPRA--QEHHEEPPGRDVL--PTQOELTAAVNPVLAKTEVQGEGBVMDLGRKVEBOE 1250
Db 1365 ENKPGQOQKTEPQKGRDIDEXOTAELELKGPEP-----KEKULD-----REDG 1409
Qy 1251 VEVHSGPNQK--AADTVYSEWVGAVACQEKESTEVQSLSEEGEM--ETDVEKEKRETP 1308
Db 1410 --QRAQPAKAKASQRTDLD-----QOTQATEPRDAQGERDSEKDKSLFLDRTP 1461
Qy 1309 EGVSEGEQETAPEHEGTGKVPILTDM-----SEERKAL-----GSIG 1350
Db 1462 BEKDRILVQEDBAPEH--SIEPTQIDRAPDRKGTDEKQEKERASEKQVLEQKQWALG 1519
Qy 1351 GSPSLPQOKAGIEVQVQSL--DTVTQTAAEVEKVI-----ETVVISSEGESPECV--- 1401
Db 1520 KEGETLDOE--ARTAEQKDETLAKDKTQOGKSFVEDKTTTSKETVLDOKSAEADSVEQ 1578

```

```

Qy 1402 -GAIL-----LPAKSSATGS-----HW-----TLQHAEDTVPLGPE---SQAESPI 1440
Db 1579 DGAALFKTRALGLEESPEPSKAREQKRYKQEDVVQGMRETSPTGEPVGGKEPVPA 1638
Qy 1441 IVPAPESSTLH-----PDLQGISASORERS-----EEDKPDAGPD--ADGKESTAI 1487
Db 1639 WEGSPQEVVYWRDRDITLQODAYWRELSCDRKVPHPHLDGGARPRYCEBERESTFLD 1698
Qy 1488 KVLAAPEPILELESKNKIVLVNIQTAVDQFARTETAPETHAY--DSGTQVPACRIDSREP 1546
Db 1699 EG-PDEGRITPLQ-----HTPRSPWTSDFQDFQEPPLQKGLVEYER- 1737
Qy 1547 NRCWTMKMDACKMHPVQPPREDLOYLT 1573
Db 1738 --WL-----AESPPGLPPEEDKLT 1755

RESULT 7
MAPA_HUMAN
ID MAPA_HUMAN STANDARD; PRT; 2805 AA.
AC P78559; Q15882; Q12973; Q9UT74; O95643;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Microtubule-associated protein 1A (MAP1A) (Proliferation-related
protein p80) [Contains: MAP1 light chain LC2].
GN MAP1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97001161; PubMed=8812494;
RA Fink J.K., Jones S.M., Esposito C., Wilkowski J.;
RT "Human microtubule-associated protein 1A (MAP1A) gene: genomic
organization, cDNA sequence, and developmental- and tissue-specific
expression.";
RL Genomics 35:577-585(1996).
RN [2]
RP SEQUENCE OF 1-1825 FROM N.A.
RX TISSUE=Ovarian carcinoma;
RA Chen Z.C., Fadiel A., Natfolin F.;
RT "Identification of a novel protein (p80) in ovarian carcinoma cells.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 78-1687 FROM N.A.
RX TISSUE=Brain;
RA Ohnani K., Rutherford T., Sakamoto H., Natfolin F.;
RT "Microtubule associated protein 1A (MAP1A) in human brain - DNA
sequence and physiological role.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 134-419 FROM N.A.
RX TISSUE=Fetal muscle;
RA Chamilikhal N., Paeturaud P., Richard I., Auffray C.,
RA Beckmann J.S.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1607-1883 FROM N.A.
RX TISSUE=Brain;
RA MEDLINE=95356255; PubMed=7629894;
RA Fukuyama R., Rapoport S.I.;
RT "Brain-specific expression of human microtubule-associated protein 1A
(MAP1A) gene and its assignment to human chromosome 15.";
RL J. Neurosci. Res. 40:820-825(1995).
CC -I- FUNCTION: Structural protein involved in the filamentous cross-
bridging between microtubules and other skeletal elements.
CC -I- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate
with MAP1A and MAP1B proteins.
CC -I- TISSUE SPECIFICITY: BRAIN.
CC -I- DOMAIN: THE BASIC REGION CONTAINING THE REPEATS MAY BE RESPONSIBLE

```

CC FOR THE BINDING OF MAP1A TO MICROTUBULES.
 CC -!- PTM: VARIOUS SERINE RESIDUES MAY BE PHOSPHORYLATED BY CAMP KINASE.
 CC -!- PTM: LC2 IS COEXRESSED WITH MAP1A. IT IS A POLYPEPTIDE GENERATED
 CC FROM MAP1A BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
 CC BOTH MAP1A AND MAP1B.
 CC -!- SIMILARITY: TO MAP1B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U38291; AAB41132.1; -;
 CC EMBL; U38292; AAB41133.1; -;
 CC DR EMBL; AF200415; AAF08305.2; -;
 CC DR EMBL; U80458; AAD00355.1; -;
 CC DR EMBL; Z47038; CAA87104.1; -;
 CC DR EMBL; U14577; AAB81362.1; -;
 CC Genew; HGNC:6835; MAP1A.
 CC MIM; 600178; -;
 CC KW Microtubules; Repeat; Phosphorylation.
 CC FT CHAIN 22490 2805 MAP1 LIGHT CHAIN LC2.
 FT DOMAIN 309 496 LYS-RICH (BASIC).
 FT 415 541 9 X 3 AA REPEATS OF K-K-[DE].
 FT REPEAT 415 417 1.
 FT REPEAT 420 422 2.
 FT REPEAT 427 429 3.
 FT REPEAT 431 433 4.
 FT REPEAT 436 438 5.
 FT REPEAT 440 442 6.
 FT REPEAT 444 446 7.
 FT REPEAT 449 451 8.
 FT REPEAT 539 541 9.
 FT CONFLICT 72 72 F -> L (IN REF. 1).
 FT CONFLICT 134 135 VV -> IP (IN REF. 4).
 FT CONFLICT 249 249 A -> G (IN REF. 4).
 FT CONFLICT 263 263 V -> A (IN REF. 4).
 FT CONFLICT 296 296 Q -> H (IN REF. 3).
 FT CONFLICT 311 311 S -> G (IN REF. 4).
 FT CONFLICT 324 324 K -> Q (IN REF. 1).
 FT CONFLICT 335 336 AK -> ST (IN REF. 1).
 FT CONFLICT 353 353 A -> S (IN REF. 1).
 FT CONFLICT 357 357 A -> S (IN REF. 1).
 FT CONFLICT 364 364 K -> Q (IN REF. 1).
 FT CONFLICT 414 419 EKKDKE -> KKKRNS (IN REF. 4).
 FT CONFLICT 424 424 K -> P (IN REF. 1).
 FT CONFLICT 426 426 E -> D (IN REF. 3).
 FT CONFLICT 431 431 K -> Q (IN REF. 3).
 FT CONFLICT 439 439 E -> D (IN REF. 3).
 FT CONFLICT 444 444 K -> R (IN REF. 3).
 FT CONFLICT 452 453 TK -> SS (IN REF. 3).
 FT CONFLICT 457 457 K -> R (IN REF. 3).
 FT CONFLICT 682 682 Q -> P (IN REF. 1).
 FT CONFLICT 1025 1025 Q -> K (IN REF. 1).
 FT CONFLICT 1303 1313 KVLPGAITSPD -> EVLTWGDHQAALN (IN REF. 3).
 FT CONFLICT 1335 1341 MISSING (IN REF. 3).
 FT CONFLICT 1368 1368 Q -> T (IN REF. 3).
 FT CONFLICT 1470 1470 A -> T (IN REF. 3).
 FT CONFLICT 1650 1650 W -> C (IN REF. 1).
 FT CONFLICT 1690 1690 A -> S (IN REF. 1).
 FT CONFLICT 1714 1714 G -> V (IN REF. 1).
 FT CONFLICT 1869 1869 E -> A (IN REF. 5).
 FT CONFLICT 1879 1883 GTPPEY -> AHSRV (IN REF. 5).
 SQ SEQUENCE 2805 AA; 306392 MW; F3ED0A3165993B2E CRC64;
 Query Match 5.2%; Score 421; DB 1; Length 2805;
 Best Local Similarity 20.8%; Pred. No. 1.5e-07;
 Matches 349; Conservative 237; Mismatches 671; Indels 420; Gaps 72;
 QY 103 VEEMAANSTAVEDIT-KDGOE-----ETSEIIIEIQIPASENNVEEMVQ---PAESQA 149

Db 216 MQKWAGNSIAKGTGIVLPNGKEAIEISVPYLTSTALVWMLPA--NPTEKIVRVLPFGNAPOQ 273
 QY 150 NDV--GFKKV--FKVGFKFTVKOKNEKSDTVQLLTVKKDEGEAEASVGAGDHQEPSV 205
 Db 274 NKILEGLEKLRLHLDLRYPVATQKLAGAVPTNLKPSIKQRADSKESLKA-----TT 327
 QY 206 ETAVGESAKSELKOSTKQGTGLKQEOSSTEIPLQASDQAAEBEADGEGEKEKEP 265
 Db 328 KTAVSLAKREVEVGEAKARSALAKELAKT-----EKKAKESSE----- 368
 QY 266 TKSPESSPSP--VNSPTTSSGKFFTHGWAGRKKTSPFKSK--EDDLTAERKKEQEA 320
 Db 369 -RPPEKPAKPERVKTESSEALK-----AERKLIKQVKGKHLKEKISKLEEKDKK 420
 QY 321 EKVDDEEKE--KTEPASEQEPAEDTDQARLSADYKVEKLEPLEDQVGDLEASSEKCAPL 378
 Db 421 KEIKKERKELKXDEGRKEEKDA-----KKEEKRDTKPELKIKSKPDLKPF 467
 QY 379 ATEVDFEKMEAHQEVVAEVHVSTVEKTEBEEQGGGAEAGVVVEGTGSLPPEKLAEPQE 438
 Db 468 TPEV-----RKTLYKAKVGRVKIDRSRAIRGE-----KELSEPOQ 504
 QY 439 VPQE--AEPAEELMKSRMCVSG-GDHTQ-LTDLSPEEKT--PKHPEGIVSEVEMSSQ 492
 Db 505 PPAQKGTVPPTISGHRELVLSSPEDLTQDFEEMKREERALLAEQDRTGLGDKPFLDTA 564
 QY 493 ER-----IKVQGGSP-----LKKLFSSSGLKKLKGKQKRGKGGGDEEPEYQH 536
 Db 565 EGGPSTAIQGTTPSPVGLQBEHVMKELVPEPBEQSGKDRGLDSCAETEE--EKDT 622
 QY 537 IHTESPESADEQKGESSASSPEPE-ETTCLEKGLPEA-----POGRAEBEGTSGEKK 590
 Db 623 WBEKQREARLPRDTEAREESEPEVKEDVIEKAELEEMEVEHPSPDEEDATKAEGFYQ 682
 QY 591 REGITPWASFKQWVTPKKVRPSPESDKHEEL-----EKVKSATLSSTDST-----V 637
 Db 683 -----KMQBPLKVTTPRSREAFGRELGLQGAPEKETSLSLSTTPAGATEHV 732
 QY 638 SEMQDEV-----KTVGEEQKPEPKRRVDTSVSWEALICVGSKKRARKASSDDEG 689
 Db 733 SVIQDETIPGYSETETQISDEIHDEPEER-----PAPRPHSTSYDLP 776
 QY 690 GPRTLGGDSHRAEASKOKAGTAVPASTQDQAGSSSP-----AGSPSEG 740
 Db 777 GPEGAG-----PFEASQPADS---AVPATS-----GKXAGTPELTPTNIVAAPLAE 823
 QY 741 EGVSTWESFKRLVTPKKSLSKLEAKEDSSVEQLSTEIPEPSRESWS-----IKKFP 795
 Db 824 EHVSSATS-----ITECDKLSSFATSVAEDQSVALTA---PQTEETGKSLLLDVTSP 876
 QY 796 GRRKKGADGKQQAATVEDSGPVEINEDDPNVAVPLSEYNAV-EREKMEAQNTLPOL 854
 Db 877 SSRTEATQGLDVPSAGTISPTSSLEEDGFKS-PCEDFSVTGSEKR----- 924
 QY 855 LGAVVYSELSKTLVHTVSAVVDGTRAVTSVEERSPS-----WISASVTEPLEHTAGRA 909
 Db 925 -----GEIIGKGL-----SGRAVEEBEETANVEMSEKLSQYGTVPFSAPGHA 969
 QY 910 MPVPE-----EVTEKDIIAEETPV-LTQTLPEGKDA--HDDMVTSEVDFTSSEATATET-- 960
 Db 970 LHPGEPALGEAEERCLSPDDSTVKMASPPSPSATHPPHQSPVEEKSEFPDFOEADS 1029
 QY 961 -SEALRTEBVTASGAETTTDMVASVSLTD-----SPDTTEATPVQVEESG--- 1007
 Db 1030 WGDTRKTPGVGKEDAAEETVKFGPEGTELEKEKVPKPSQAQAPVNIDEGLTCTTQ 1089
 QY 1008 -----VLDEEERQTAILOAVA-----DKVKEESQVPATQVORTGSKALEKVE 1053
 Db 1090 LLPAQDKAIVFEIMEAGEPTGPILGAEALPGGLRTLPQEPGPKQKDEVLRYPDRSL--- 1146
 QY 1054 EVEEESVLA-----SEKEDVMKPGVQAGAHLAQGSQETQATPESLEVPEVTD 1106

```

Db 1147 --PEBAESLSVSVSP8PTANOEPFPPKPC--GLTEQYIHKORWREVPEDTQSLSE 1202
Qy 1107 VDHVATCCQVIKIQQLMEQAVAPESSETLTDTSETN----- 1140
Db 1203 SPSKETSILVSSKQL-----SPESLCTLTQFGEINLCKEEMGHLMQADTSHHTAPMSVPE 1257
Qy 1141 ---GSTPLADBDTADGTQODETIDSDSKATPAVQSV-----TEENA 1182
Db 1258 PHAATASPTDGTTRYSAQTDITDSDLSRKSPASSSHSTPSGNGKYLPGLAITSPEHITL 1317
Qy 1183 TQAO-----KEEPTLP-----NNVP-----AOEHEGEPGRDYLEPTQELTAA 1221
Db 1318 TDDSSSPSKPESLPGFALMEDIAIKWEDKVPGLKDRTSSEQKEPEPDEVLOQKDTLHK 1377
Qy 1222 AVPLAKTEVGEQGEVDWLDEKVKSEQEVFVHSGPNSQKADVTYDSEWGVAGCQEK 1281
Db 1378 EVVEPMDTAIYQKDEALHYKNEAVKQODKALBQK-----RDLBQKDTALEQKDK 1427
Qy 1282 STEVQSLSEEGEMETDVKEKRETPQEOVSEGEETAPEHEGTYGKPVLTLPMPSE 1341
Db 1428 ALEPKDKDE---EKDKALEQDKIPEBKDKALEQKDTALEK---DKALEPKDKDEBQ 1480
Qy 1342 RKGALGSLGSPSLPDPDAGCEVVOGSLDTVTQ--TAAVEKYIEFVVISSETGESPE 1399
Db 1481 KRVULEQKKEIE--EKDKA--LDQKRVSEVHKAPEDTYAEMKDRLE-----QTDKAP 1531
Qy 1400 CVGAHLLEPAEKSSAT-----GHWTLQHAEDTVPLGPESQAESIPITVPAPESTLA 1451
Db 1532 --QKIQAGQKQKVKSEKQKQALQKQKQWALGQKDEALQKQALEHNQ--TQEGESLVO 1586
Qy 1452 PDLQETIASQGRSEEREKPPADGPDADGKESTAIKTV-----LKAPELLEIESKS 1503
Db 1587 ED-----KTRKPKMLEKSPKPVKAMEKLELLEKTKALGLESLVOEGRAREQEEKY 1640
Qy 1504 NKIVLVNICTAVDQFARTETAPETHAVDSOTQVPACRLDSRE--PNCWTGKMKDAKK 1559
Db 1641 WR-----GQDVVOEWQETSPTRPEEPAGEKELAPAMEDTSPEDNRYMNGREVDAL 1692

```

RESULT 8
PGCV_CHICK STANDARD; PRT; 3562 AA.

```

AC Q90953; Q90945; (Created)
BT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Version core protein precursor (Large fibroblast proteoglycan)
GN (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
OS CP962.
OC Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RX STRAIN=White Leghorn; TISSUE=Limb bud;
RX MEDLINE=93300846; PubMed=8314802;
RA Shimomura T., Nishida Y., Ito K., Kimata K.;
RT "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
RT expressed during chondrogenesis in chick limb buds. Alternative
RT spliced multiforms of PG-M and their relationships to versican."
RL J. Biol. Chem. 268:14461-14469(1993).
CC -I- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -I- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -I- ALTERNATIVE PRODUCTS: At least 2 isoforms; V0 (shown here) and
CC VI are produced by alternative splicing.
CC -I- TISSUE SPECIFICITY: Prechondrogenic condensation area of
CC developing limb buds.
CC -I- DEVELOPMENTAL STAGE: Disappears after the cartilage development

```

```

CC (By similarity)
CC -I- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -I- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60226; CAA42787.1; -.
CC EMBL; D13542; BAA02742.1; -.
CC HSSP; P00740; IEDM.
CC InterPro; IPR000152; Asx hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00059; lectin_C; 1.
CC Pfam; PF00084; sushi_1.
CC Pfam; PF00193; Xlink; 2.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00001; EGF_like; 1.
CC SMART; SM00409; Ig; 1.
CC SMART; SM00445; Link; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS01241; Link; 2.
CC PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
CC PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
CC Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
CC Hyaluronic acid; Alternative splicing.
CC SIGNAL 1 26
CC CHAIN 27 3562
CC DOMAIN 37 136
CC DOMAIN 166 243
CC DOMAIN 264 345
CC DOMAIN 3254 3290
CC DOMAIN 3292 3328
CC DOMAIN 3341 3455
CC DOMAIN 3460 3518
CC DISULFID 44 129
CC DISULFID 171 242
CC DISULFID 195 216
CC DISULFID 269 344
CC DISULFID 293 314
CC DISULFID 3258 3269
CC DISULFID 3263 3278
CC DISULFID 3280 3289
CC DISULFID 3296 3307
CC DISULFID 3301 3316
CC DISULFID 3318 3327
CC DISULFID 3334 3345
CC DISULFID 3362 3454
CC DISULFID 3430 3446
CC DISULFID 3461 3504

```

FT DISULFID 3490 3517 BY SIMILARITY
FT CARBOHYD 163 235 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 235 329 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 329 529 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 529 709 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 709 948 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 948 1409 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1409 1479 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1479 1523 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1523 1530 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1530 1625 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1625 1751 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1751 1988 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 1988 2088 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2088 2089 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2089 2507 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2507 2642 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2642 2679 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2679 2748 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2748 2762 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 2762 3069 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 3069 3194 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 3194 3232 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 3232 3545 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT CARBOHYD 3545 485 N-LINKED (GLCNAC. .) (POTENTIAL) .
FT VARSPLIC 485 1411 MISSING (IN ISOFORM VI) .
SQ SEQUENCE 3562 AA; 388078 MW; 95C566E88C1602D2 CRC64;

Query Match 5.1%; Score 411.5; DB 1; Length 3562;
Best Local Similarity 19.7%; Pred. No. 4e-07;
Matches 397; Conservative 269; Mismatches 698; Indels 649; Gaps 94;

QY 49 TKLPQ--KNQLSGVNGVAB-QGDVHVQ-----EENQEQEBEWVDVQGRES 96
DB 433 TLLPQTVDGEISPYDTLGRTEYDVSPLRTESTSAALVEHTYSEALSSE--QGRSES 489
QY 97 VREKDRVEANAANSTAV-----EDITK-----DQEQETSEIIEIPASENNVEE 140
DB 490 TEDAFTSVVFQDSTAVAKSSTGSWEDIEGTQKHGDQNTQIEVGPVMTATDSLVA 549
QY 141 MVQPAESQANDVGFKKFVGFVKFT---VKDKNEKSD---TV-----OLLTVKXD--- 186
DB 550 SORLPTGSSVSLTKENLVLSHSTKEPTKKSNEAKSKLLTVTVIPKALFTDQDYLTT 609
QY 187 EGGAE-----ASVAGAGHQEPV--ETAVGESAKSELKQSTEQEGT--L 230
DB 610 GGEGRSMYTMPDRVSGVALVSPESDVPVAVSETLMDELAVTTG---QSSADESTPFI 666
QY 231 KQOSSTEIPLQESDQAAEEBAKD----- 255
DB 667 KFSSTATELDNEA-SAEGRSREDLKDVHLTTSSGIPVSVFTLTANETGSEVTALSESTSAP 725
QY 256 -----EGE---EKQEKPTKSPESPSPVN-----SETTSSFK 285
DB 726 QKFEEGITSVLHSSQTEGSAILEKQB--TKPEMSIIDAKVLVYITVVVPASVTAGSEG 783
QY 286 KP-----FTHG-----KTSFKK-----SKEDDLETAE----- 313
DB 784 RFGSEKFTHTPPVSGMWLQDQVYVMTESHTSKRIELOTEDDISGMEPTSSPGQIIEY 843
QY 314 -KKEQEAKEVDREK--EKTPEASEQEPADTDQAR-----LSADYKVV--E 357
DB 844 TKHLGAPVSAVTDETKTSMTAETESDEEVVSADFDQKTGTTTEVFHTSSSLDLKFTLSK 903
QY 358 LPLEDQVGDLEASSEKCAPLAT-----EVFDEKMEAHQEVVAHVHVSVE----- 403
DB 904 IPEDESSATVKFSFSSSGTVLPATVATVLEVTHEADETSGYVNLMTFTSPGEQKATE 963
QY 404 -----KTEEEQGGGGE-----AEGGVVVEGTGESLPPEKLAEPQOEVAEPAEELMKS 452
DB 964 KSPATSAEDEVSTGTELSKYTMTEGGQISSVT--SAEKESVAALQOE--REQPSVGLPET 1019
QY 453 REMCVSGGDHTQTLTDLSPBEETLPKHPEGIVSEVEMLSQERI----- 495

DB 1020 KE-----PFRFTDVTEIETTVPOR--EGDTSLVPTVVGSEDI GEMQVTDHTSFDSIIHT 1071
QY 496 --KVQSPPLKLFSSSGLKSKGKQKRGGG--GDDEEPGEYQHHT-----ESPESA 545
DB 1072 EATVSTKASEVFP--KELSTKDQDRELGTAMGSTLPTVSQMHEQKTTAGFESPQTT 1127
QY 546 DQKQGESSASSPEEPEETTCLE-----KGPLEAPQ-----DGEAREGITSDE 588
DB 1128 TOEKIDEMGSAYDEMYPATELSVPALMLTEYQVGSVPVETSTRSLHLTGTPKAETADQE 1187
QY 589 KK-----REGITPWASFKMVTPKKRVRRPSE----- 615
DB 1188 EKITEAVPTFGTOAKVYESKGTTRTREDRDVGSWNSVLPHTMLSSPSTAGSISLLTIG 1247
QY 616 -----SDKEELEKVKSATLS--TDST--VSEMO-----DEVKTVGEEK-- 652
DB 1248 ASPSQTPGSGIGISEELEEVKVPFSSRATDKTIVISDLTSSISAVDKIQTSAKSPFVS 1307
QY 653 -----PEEPKRVDTSVSWEALICVSSKKRARKASSDDEGGPRTLGGDSHRAEAS 705
DB 1308 SKSPRIIPE-----DEEVTSSDIIVIDESISPS-KASAEDDLGT-----KMVEPE 1352
QY 706 KDKAECTDA-----VPASTQEQDQAGSSSPEPAGSPSEGEVSTW----- 746
DB 1353 IDKEYFTSTATAVARPTAPPTVMEATEALQOEVSPTSHPDSGTDIRLYVITQINDTD 1412
QY 747 -----ESFKRLVTPRKSKLEEKARDSSVEQLSTE-IPSRSEESVWSIKKFIPIGRR 798
DB 1413 HPWNEFLDLFSRHILPHAVDEHTD--AESAQTEPTCSDSVQDSSE--YIILDPFP-- 1465
QY 799 KKRADKQKQOATVDSGPGVINEEDPNVAVPLSEYNAVERKMEQAQNTLPQL---- 854
DB 1466 -NFMDFEEEDCE-----NTTDVTTT--PALQFINGKQVTSAPKSTKAEARSQ 1514
QY 855 LGAVVYSELSKTLVHTVSVAVIDGTGTRAVTSVEERSPSWIS--VTEPLEHTAGAM-P 911
DB 1515 IESVAHKNVTFSQINETNTFIISETEAGTQWPSKAGEVMGAFVETQP--TADVAMLE 1571
QY 912 PV-----EYVTEKIIAETPVLVTQTLPEGKDA---HDDMVTSEVDFTSEAVTATSEAL 964
DB 1572 PVYSGSEVTTDKYLEITSVVEQSPKKNKVTMMHG----- 1608
QY 965 RTEEVTASGAEETDMVSAVQLTDSPTTTEATPVQEVESGLVDTEEBEERQTQAILQA 1024
DB 1609 -----TEESSTKDTKNLLITNE--SSGSGSTESOLSRSVFTTEILTMSSHED----- 1653
QY 1025 VADKVEESQVATQTVORTGSKALEKVEEVEDSEVLASEKEKDVMPKGPVQEAHEL 1084
DB 1654 -SEKISHTTSVPTILSVERS---AVTAAPSADSDTATVGID-VKDLIPKG----- 1698
QY 1085 AQSGETGOATPESLEVPVETADVH-----VATCQVIKLQQLMEQAVAPESSETLTDSE 1138
DB 1699 -----GTATPGNYKSTIKLDAEPFENPEATSHHTKPDMTASSFVLESGDVEENS 1752
QY 1139 TNGSTPLADSTADGTQDQDETIDSDQSKATAAVROSQVTEEBEATAQKEEPTLPNNVPA 1198
DB 1753 T-----LAGAMTTE-TAVAETLSVQDT-----SIGSGTVLPTISVITSEITPALPG-- 1798
QY 1199 QBEHGEERDVLPTQOELTAAAVPLAKTEVGOGEVDMLDGEKVEEQEVFVHSGPN 1258
DB 1799 -----GTRILYSTFQDSEATVNFVSELIMEQVY---GSSVATEKKV-----ED 1841
QY 1259 SOKAADVTYDSEVMGVAGQOEKE-----STEVQSLSLE-----EGEMETDVE 1300
DB 1842 EKEVQTTVSSGEIETTDKAGKSELDEFSGTNEVTVSQEPTPLREIVPTGTMHSEIK 1901
QY 1301 KKKRE--TKPEQVSEGEQETAAPHEGYKPVLTLDMPSSERGA-LGSLGSGSLPD 1357
DB 1902 KVTATPFLREKLFINEGSAEAPADLFAGSPTRKVVSTDSPTDSGSDIDVITESATLFS 1961
QY 1358 QKAGCIEVQ-----VQSLDTT-----VTQTAEAV 1382

Db 1962 VPSRVIETQTVKHEGNIIVIVSVLSKNTTTEYEHIGTGVPVTSVSGSDGJTERSEVA 2021
 Qy 1383 EXVIETVVISFEGESP--ECVGAHLIPAKSGATGGMWTLQHAEDTVPLGPSQAEISPI 1440
 Db 2022 IEMSEVFEFTENQGEPTQAVPTVTPSDIKSRIG-----SRREVTSHVTPVTRTDLET 2076
 Qy 1441 I-VTAPAEPTLHPD-----LOGEISASQRESEEDKPDAGDADGKSTALEKTKAE 1493
 Db 2077 AEVTSPEVSVNNSTLDVTWHTITRAVAESTESKKG-----GSFASVSGKILMTE 2129
 Qy 1494 ---PELLEESKSNKIVLN-----VIGTAVDQFARTETAPE----- 1526
 Db 2130 HOSGELKNDSSSTTKLMSGPTREKLGHSPFDGSGAEFTTESFTASVSPTKPEP 2189
 Qy 1527 -----THAYDSQTVPAACRLDSREPNRCWK-----MKDAMKE- 1560
 Db 2190 QEOYGRKTVSMPSAVVHAH-----TAEPNELVTSTEDHTISLQTVTDTMEBK 2237
 Qy 1561 ----PVQPREDLQVLTVLEAMA-QPRKCLR 1587
 Db 2238 ANNELTVTSFATNLPLESDVHSMEDPREILPK 2270

RESULT 9
 PLE1 RAT STANDARD, PRT; 4687 AA.
 AC P30427; 008879; 008880; 008881;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Plectin 1 (PLTN) (PCN).
 GN Plectin.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=glial tumor;
 RX MEDLINE=91268156; PubMed=2050743;
 RA Wiche G., Becker B., Luber K., Weitzer G., Castanon M.J.,
 RA Hauptmann R., Stratawa C., Stewart M.;
 RT "Cloning and sequencing of rat plectin indicates a 466-kD polypeptide
 chain with a three-domain structure based on a central alpha-helical
 coiled coil.";
 RT J. Cell Biol. 114:83-99(1991).
 RN [2]
 RP REVISIONS.
 RC TISSUE=glial tumor;
 RX MEDLINE=96210632; PubMed=8633055;
 RA Liu C.-G., Maercker C., Castanon M.J., Hauptmann R., Wiche G.;
 RT "Human plectin: organization of the gene, sequence analysis, and
 RT chromosome localization (8q24).";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4278-4283(1996).
 RN [3]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS 2; 3; 4), AND TISSUE SPECIFICITY.
 RC TISSUE=glial tumor;
 RX MEDLINE=97321050; PubMed=9177781;
 RA Elliott C.E., Becker B., Oehler S., Castanon M.J., Hauptmann R.,
 RA Wiche G.;
 RT "Plectin transcript diversity: identification and tissue distribution
 RT of variants with distinct first coding exons and rodless isoforms.";
 RL Genomics 42:115-125(1997).
 CC -I- FUNCTION: INTERLINKS INTERMEDIATE FILAMENTS WITH MICROTUBULES AND
 CC MICROFILAMENTS AND ANCHORS INTERMEDIATE FILAMENTS TO DESMOSES OR
 CC HEMIDESMOSES. MAY BE INVOLVED NOT ONLY IN THE CROSSLINKING AND
 CC STABILIZATION OF CYTOSKELETAL INTERMEDIATE FILAMENTS NETWORK, BUT
 CC ALSO IN THE REGULATION OF THEIR DYNAMICS.
 CC -I- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
 CC -I- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1 (SHOWN HERE), 2, 3 AND 4; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSION IN
 CC SKELETAL MUSCLE AND LOWEST IN THYMUS.

CC -I- DOMAIN: THE N-TERMINUS INTERACTS WITH ACTIN, THE C-TERMINUS WITH
 CC VIMENTIN, DESMIN, GFAP, CYTOKERATINS, LAMIN B, WHEREAS BOTH THE N-
 CC AND THE C-TERMINUS CAN BIND INTEGRIN BETA-4.
 CC -I- PTM: PHOSPHORYLATED BY CDC2; REGULATES DISSOCIATION FROM
 CC INTERMEDIATE FILAMENTS DURING MITOSIS (BY SIMILARITY).
 CC -I- SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.
 CC -I- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -I- SIMILARITY: CONTAINS 3 PLECTIN REPEATS.
 CC -I- SIMILARITY: CONTAINS 4 SPECTRIN REPEATS.
 CC -I- SIMILARITY: BELONGS TO THE PLEKIN OR CYTOLINKER FAMILY.
 CC -I- CAUTION: ISOFORM 4 IS A FRAGMENT AT THE N-TERMINUS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; X59601; CA442169.1; -;
 CC EMBL; U96274; AAC53209.1; -;
 CC EMBL; U96275; AAC53210.1; -;
 CC EMBL; U96276; AAC53211.1; -;
 CC PIR; A39638; A39638.
 CC PIR; S21876; S21876.
 CC HSSP; Q01082; 1BRK.
 CC InterPro; IPR001589; Actbind actin.
 CC InterPro; IPR001715; Calponin-like.
 CC InterPro; IPR001101; Plectin repeat.
 CC InterPro; IPR005326; S10_plectin.
 CC InterPro; IPR002017; Spectrin.
 CC Pfam; PF00307; CH; 2.
 CC Pfam; PF00681; Plectin; 21.
 CC Pfam; PF03501; S10_plectin; 1.
 CC SMART; SM00033; CH; 2.
 CC SMART; SM00250; PLEC; 33.
 CC SMART; SM00150; SPEC; 4.
 CC PROSITE; PS00019; ACTININ_1; FALSE NEG.
 CC PROSITE; PS00020; ACTININ_2; FALSE NEG.
 CC PROSITE; PS50021; CH; 2.
 CC Coiled coil; Repeat; Structural protein; Cytoskeleton; Actin-binding;
 CC KW Phosphorylation; Alternating splicing.
 CC FT DOMAIN 1 1473 GLOBULAR 1.
 CC FT 1474 2758 CENTRAL FIBROUS ROD DOMAIN.
 CC FT DOMAIN 2 2759 4687 GLOBULAR 2.
 CC FT 181 406 ACTIN-BINDING.
 CC FT DOMAIN 3 185 288 CH 1.
 CC FT DOMAIN 4 301 403 CH 2.
 CC FT REPEAT 648 722 SPECTRIN 1.
 CC FT REPEAT 743 827 SPECTRIN 2.
 CC FT REPEAT 840 933 SPECTRIN 3.
 CC FT REPEAT 1318 1418 SPECTRIN 4.
 CC FT DOMAIN 1472 1692 COILED COIL (POTENTIAL).
 CC FT 1724 2760 COILED COIL (POTENTIAL).
 CC FT DOMAIN 2761 2828 PLECTIN 1.
 CC FT REPEAT 2829 2866 PLECTIN 2.
 CC FT REPEAT 2867 2904 PLECTIN 3.
 CC FT REPEAT 2905 2942 PLECTIN 4.
 CC FT REPEAT 2943 2980 PLECTIN 5.
 CC FT REPEAT 2984 3018 PLECTIN 6.
 CC FT REPEAT 3119 3156 PLECTIN 7.
 CC FT REPEAT 3157 3194 PLECTIN 8.
 CC FT REPEAT 3195 3232 PLECTIN 9.
 CC FT REPEAT 3233 3270 PLECTIN 10.
 CC FT REPEAT 3271 3308 PLECTIN 11.
 CC FT REPEAT 3311 3346 PLECTIN 12.
 CC FT REPEAT 3488 3525 PLECTIN 13.
 CC FT REPEAT 3526 3563 PLECTIN 14.
 CC FT REPEAT 3564 3601 PLECTIN 15.
 CC FT REPEAT 3602 3639 PLECTIN 16.
 CC FT REPEAT 3643 3677 PLECTIN 17.
 CC FT REPEAT 3823 3860 PLECTIN 18.

DB 2805 SNEELORLTQGHHTVAELTQREDEVRYHKLKSGSSIAGLLKPTNEKLSVTALQROLSP- 2863

OY 1400 CVCAGHLLEPAKSSATG-----GHWTLOHAEDEVPLGPE 1432

DB 2864 --GTALLILEAQASGFLDPVARNRLTYNEAVKESGVPE 2902

RESULT 10

PLEI CRIGR STANDARD; PRT; 4473 AA.

AC 090155;

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Plectin 1 (PLTN) (PCN) (300-kDa intermediate filament-associated protein) (IFAP300) (Fragment).

GN Plectin.

OS Cricetulus griseus (Chinese hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus.

OX NCBI TaxID=10029;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20334248; PubMed=10873583;

RA Clubb B.H., Chou Y.-H., Herrmann H., Svikkina T.M., Borisy G.G., Goldman R.D.;

RA "The 300-kDa intermediate filament-associated protein (IFAP300) is a hamster plectin ortholog.";

RT Biochem. Biophys. Res. Commun. 273:183-187(2000).

RL [2]

RP PHOSPHORYLATION.

RX MEDLINE=96215219; PubMed=8626512;

RA Malecz N., Foisner R., Stadler C., Miche G.;

RA "Identification of plectin as a substrate of p34cdc2 kinase and mapping of a single phosphorylation site.";

RT J. Biol. Chem. 271:8203-8208(1996).

RL -1- FUNCTION: INTERLINKS INTERMEDIATE FILAMENTS WITH MICROTUBULES AND MICROFILAMENTS. MAY BE INVOLVED NOT ONLY IN THE CROSSLINKING AND STABILIZATION OF CYTOSKELETAL INTERMEDIATE FILAMENTS NETWORK, BUT ALSO IN THE REGULATION OF THEIR DYNAMICS.

CC -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER.

CC -1- DOMAIN: THE N-TERMINUS INTERACTS WITH ACTIN. THE C-TERMINUS WITH VIMENTIN, DESMIN, GAP, CYTOKERATINS, LAMIN B; WHEREAS BOTH THE N- AND THE C-TERMINUS CAN BIND INTEGRIN BETA-4.

CC -1- PTM: PHOSPHORYLATED BY CDC2; REGULATES DISSOCIATION FROM INTERMEDIATE FILAMENTS DURING MITOSIS.

CC -1- SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.

CC -1- SIMILARITY: CONTAINS 2 CALPONTIN-HOMOLOGY (CH) DOMAINS.

CC -1- SIMILARITY: CONTAINS 32 PLECTIN REPEATS.

CC -1- SIMILARITY: CONTAINS 4 SPECTRIN REPEATS.

CC -1- SIMILARITY: BELONGS TO THE PLAKIN OR CYTOLINKER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF260753; AAF70372.1; -.

DR HSSP; O01082; 1BRK.

DR InterPro; IPR001589; Actbind actinin.

DR InterPro; IPR001715; Calponin-like.

DR InterPro; IPR001101; Plectin_repeat.

DR Pfam; PF00307; CH; 2.

DR Pfam; PF00681; Plectin; 20.

DR SMART; SM00033; CH; 2.

DR SMART; SM00250; PLEC; 32.

DR SMART; SM00150; SPEC; 4.

DR PROSITE; PS00019; ACTININ_1; PARTIAL.

DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.

DR PROSITE; PS50021; CH; 2.

KM Coiled coil; Repeat; Structural protein; Cytoskeleton; Actin-binding; Phosphorylation.

FT NON TER 1 1

FT DOMAIN <1 1259

FT DOMAIN 1260 2544

FT DOMAIN 2545 4473

FT DOMAIN <1 192

FT DOMAIN <1 74

FT DOMAIN 87 189

FT DOMAIN 449 508

FT REPEAT 529 613

FT REPEAT 626 719

FT REPEAT 1104 1204

FT DOMAIN 1258 2548

FT REPEAT 2615 2652

FT REPEAT 2653 2690

FT REPEAT 2691 2728

FT REPEAT 2729 2766

FT REPEAT 2770 2804

FT REPEAT 2905 2942

FT REPEAT 2943 2980

FT REPEAT 3019 3056

FT REPEAT 3057 3094

FT REPEAT 3274 3311

FT REPEAT 3312 3349

FT REPEAT 3350 3387

FT REPEAT 3388 3425

FT REPEAT 3429 3463

FT REPEAT 3609 3646

FT REPEAT 3647 3684

FT REPEAT 3685 3722

FT REPEAT 3723 3760

FT REPEAT 3764 3797

FT REPEAT 3800 3834

FT REPEAT 3852 3889

FT REPEAT 3927 3964

FT REPEAT 3965 4003

FT REPEAT 4007 4041

FT REPEAT 4043 4094

FT REPEAT 4197 4234

FT REPEAT 4235 4272

FT REPEAT 4273 4310

FT REPEAT 4311 4348

FT REPEAT 4349 4386

FT DOMAIN 4039 4089

FT DOMAIN 4414 4429

FT MOD RES 4328 4328

SO SEQUENCE 4473 AA; 509015 MW; E144615D361E3484 CRC64;

Query Match 5.0%; Score 402.5; DB 1; Length 4473;

Best Local Similarity 19.2%; Pred. No. 1e-06;

Matches 343; Conservative 280; Mismatches 774; Indels 389; Gaps 64;

OY 8 EGRSPRQPRAGSP-----TSELVLSGHPRAEASGAPPADADPATKLPK-----NQO 57

DB 1273 EERLAEQQAERERERLAEAGALEEKOROLAEEHAQAKQAELIE-AQELQRRQGEVARRE 1331

OY 58 LSSVNGVAEQGDVHVOENQ---EGQEEVDE---DVQRESEDEVKDRVEEMAANS 110

DB 1332 EAAVN--AQQQRKRSIQEELQHROSSEAEIQAKAOVLAESRRIRIEELRVVALQLET 1389

OY 111 TVAVEDITYDGGQSETSSEI-IEQIPASENVVEVQVPAESQANDVGFKVFKFVGKFTYK- 168

DB 1390 TERQSGAAGEGELQALRAEAEEAQAQQAQAEERLRQVQDESQRKQAEEALRLRYVA 1449

OY 169 --KDKNEKSDTYQLLTIVKKDEGEAGASVAGADHQ-----PPSVETANGESAKSES 217

Db 1450 QAEAAQKQALQALELRLQAEAEARRLQAQARARQVQVALETAQRAEAEVELQSKRA 1509
Qy 218 ELKOSTEQBGLTKQEO--SSTEIPLQAE---SDQAEAEAEKDEGEKQEKPTKSPSPS 273
Db 1510 SFAEKTAQLERTLQEEHVTVTQLREKAERAAQQAQAEARAEARELERWQLKANEALR 1569
Qy 274 SPVNSEITTSFKFFTHGAGWRKKTSPKSKEDDLTAEKREKQAEKVDSEKTEP 333
Db 1570 LRQAEVAQOQSL-----AQADAQKQKEAEAREARRRGAEEQAVRQRELAQOE- 1619
Qy 334 ASBEOEPAEDTDQARLSADYKVELPLEDOVGDLQAS--SEKCAPL---ATEVFDEKMEA 389
Db 1620 LEQORQLAEGTAQORLAABEQELRLRAETEGQEQORQLLEELARLQREATAATHKQOEL 1679
Qy 390 HQEVV-----AEHVSTVTEKTEBEGGGGE-----AEGVVVEGTGESL 428
Db 1680 EAEALKVRAEMEVLLASKARAEBSRSTSEKSKORLEAEADRPRELAEEAARALAAEEA 1739
Qy 429 PPEK-LAPQOEVPQEAAPAEELMKSRMCVGGDHTOLTDLSPEEKTLPHKPEGIVSEVE 487
Db 1740 KRQROLAEDAARQRAE-AERVLTEKLAAYS-----EATRLKTEAEATLKEKEAENELR 1793
Qy 488 MLSSQERIKVQSGPLKPLFSSGGLKLGKKKQKGGGDEBEGVQHIHTSPESADE 547
Db 1794 RLAEDEAFQ-----RRLEQOALHKADIERLAQ-----LRKASELER 1834
Qy 548 QKG-ESSASSPEPEPTTCLKGPLEAPQDGEAE-----EGTTSDEKKREGITPW 597
Db 1835 QKGLVEDTLRQRRQVEEILALKVSPKAAAGKAELELGRIRSSAEDTMRSEKQAEQE 1894
Qy 598 ASFKQMTVPKVRPRPSDK-----EELKVKQATLSSTDSTVSEMDQEVKTVGEE 650
Db 1895 AARQROLAEEBQRRREABERVQKSLAAEEBAARQKAALEEVERLKAV-EEARRLRER 1953
Qy 651 QKPEEPK-----RRVDTSVSWALICVGSKKRKAASSDDGEGPRTLGGDSHR 700
Db 1954 ABESARQLQAEAAQKRLQAEKAHAFV--QORREELQOTLQOQSMLERLRGAE 2011
Qy 701 AEFASKD-KEAGTDVAVPASTQEQDQAQGS-----SSPEPAGSPSEGGVSTWBSFKRLVT 754
Db 2012 ARRAAEAEAEQAEAREAAQSRKQVEAEERLQSGAEQAEQAAQAAAA-----2061
Qy 755 PRKSKSLEKEADSDSVQSLSTEIPSRSESVISIKFIPGRKRRKADQKQEQATVDS 814
Db 2062 -EKLKEAEQAEARRAQAEQAALQKQAAQDAEMEKHKFAEQTLRQAEVQELTTLR-- 2118
Qy 815 GPVEINDDPNVPAVPLSEVNAVERKME-AQNTLQOLLGAVVYS-EELSKTLVHTV 872
Db 2119 --LQLEETD-HQKSILD-EELQRLKAEVTEAARQSRQVEBELFSVRVOMBELGK-----2168
Qy 873 SVAVIDGTRAVTSVEERSPSWISASVTEP-LEHTAGAMPPEVTEVKDIIABETPVLQ 931
Db 2169 -----LKARIEAENRALILRDKNTQRFLEEEAEKMKQVAAEAAEALSLVAQAEARLQ 2221
Qy 932 -----TLPEGKADHDMVTSEVDFTSVAVTATETSEALRTEBEVTASGAEBTTMVSQV 987
Db 2222 LABEDLAQORALAEKMLKERMQAVQAEATRLKAEAEELQOQKELAQEQARRLQE-----2274
Qy 988 LTSDSPDTEATPQEVESGLVDTEEEERQTOAILQAVDKV-----EES 1033
Db 2275 -----DKEQNAQQLVEETQGFORTLEVRQRLQEMSAERLKLRLMAEMSGRAQARAEDA 2329
Qy 1034 Q-----VPATQT--VORTGS-----KALEKVEBEVDS-- 1060
Db 2330 QRFKQAEIEGKHLRTELATQEKVTLVQTLQEQSDHDAERLRAIAELEREKELK 2389
Qy 1061 -----VLASEKEDVMPKGPVQAGAEHLAQSGSETQCATPESLEVPVADVHVATQC 1114
Db 2390 QEAKLLQKSEEMTVQEQQLQTOALQKSFLEKSDLSLLQRRFIEQEKAKLEQLQFDE 2449
Qy 1115 VIKLOQLME-----QAVAPSETLTDSETNGSTPLADSDTADGTQDE-----TIDSQDS 1165
Db 2450 VAKAQQLREBQQRQOQMEQEKQELVASMEBAERRRQCEAEAEVRRKQOELQHLQELQOQ 2509

Qy 1166 KATAAVRQSQVTE-----EEATAQKEBPTLPPNNVPAQEEHGEHPGRD 1209
Db 2510 EKLLAENQRLRQLRLEEBEHRALAHSEETAATAQAAAAKALPN-----GRD 2557
Qy 1210 VLEPTQOEL-----TAAAVPVLAKEVGOEGEVDLMDGKVKKEQEVFVHSGNSOK 1261
Db 2558 ALDGPMSVEPEHAFFGLRQKVPALQREAG-----ILSAEELQLEOQHTTVAELSQR 2611
Qy 1262 AADVTYDSEVMGVAGQCEKESTEVQS-----LSLEEGEMET-----DVEKE 1302
Db 2612 EDVRYQLQGRSSIAGLLAPTEDEKLSVYTALQRLQLSLPQTALILLAQAASGFLDPVRN 2671
Qy 1303 KRETKEQVSEEGEQETAAPHEHGTYGKPVLTLDMPSSR---GKALGSLGGSPSLPDQD 1359
Db 2672 RRLTVNE-----PVKEGVVG-PELHHKLLSAERAVTGYKDPYTGQISLFOAM 2718
Qy 1360 KACCEVQVQSLDVTVTQTAEAKEVKIETV-----VISETG-- 1395
Db 2719 KK---DLLVRDHAIRLLEAQIATGGIITVHSHRVPDVAYQRGYFDEEMSRILADPGDD 2775
Qy 1396 -----ESPECVGAHLLPAEKSSATGCH--WTLQHAEDTVPLG 1430
Db 2776 TKGFFDPNTHENTYLLQLLERCVEDPE-TGLHLLPLTDKAAGGELVYDTTEARDVF--- 2831
Qy 1431 PESQAESIPITVPAPESTLHPDQGE-ISASQSRSESEEDKPDAGPDADGKSTAIEKV 1489
Db 2832 -EKATVSAPFGKQGTVTIWEIINSEYFTAQORRDLQFRGT-----RITVEKI 2881
Qy 1490 LKAEPILLESKSNKVLNVQTAVDQPARTETAPETAYDSOTQ 1535
Db 2882 IKIVITVBEQERKQGLCFEGLRALVPAAELLESVISHELYQQLQ 2927
RESULT 11
PLE1_HUMAN STANDARD; PRT; 4684 AA.
AC Q15149; Q16640; Q15148;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HD1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=96210632; PubMed=8633055;
RA Liu C.-G., Maercker C., Castranon M.J., Hauptmann R., Wiche G.;
RT "Human plectin: organization of the gene, sequence analysis, and
RT chromosome localization (8q24).";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4278-4283 (1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND DISEASE.
RX MEDLINE=96312447; PubMed=8698233;
RA McLean W.H.I., Pulkkinen L., Smith F.J.D., Rugg E.L., Lane E.B.,
RA Bullrich F., Burgeson R.B., Amano S., Hudson D.L., Owarike K.,
RA McGrath J.A., McMillan J.R., Eady R.A.J., Leigh I.M., Christiano A.M.,
RA Uitto J.;
RT "Loss of plectin causes epidermolysis bullosa with muscular dystrophy:
RT cDNA cloning and genomic organization.";
RL Genes Dev. 10:1724-1735 (1996).
RN [3]
RP VARIANT MD-EBS 1003-GLN--ALA-1005 DEL.
RX MEDLINE=97049959; PubMed=8894687;
RA Pulkkinen L., Smith F.J.D., Shimizu H., Murata S., Yaeita H.,
RA Hachisuka H., Niehikawa T., McLean W.H.I., Uitto J.;
RT "Homozygous deletion mutations in the plectin gene (PLEC1) in patients
RT with epidermolysis bullosa simplex associated with late-onset
RT muscular dystrophy.";

FT	CONFLICT	1789	1789	L -> A (IN REF. 2).
FT	CONFLICT	1910	1910	K -> R (IN REF. 2).
FT	CONFLICT	2154	2154	K -> N (IN REF. 2).
FT	CONFLICT	2160	2160	S -> R (IN REF. 2).
FT	CONFLICT	2215	2215	R -> Q (IN REF. 2).
FT	CONFLICT	2244	2244	S -> A (IN REF. 2).
FT	CONFLICT	3027	3027	K -> E (IN REF. 2).
FT	CONFLICT	3310	3310	E -> A (IN REF. 2).
FT	CONFLICT	3361	3361	F -> L (IN REF. 2).
FT	CONFLICT	3408	3408	F -> L (IN REF. 2).
FT	CONFLICT	3447	3447	G -> A (IN REF. 2).
FT	CONFLICT	3531	3531	G -> A (IN REF. 2).
FT	CONFLICT	3580	3580	R -> S (IN REF. 2).
FT	CONFLICT	3589	3589	K -> Q (IN REF. 2).
FT	CONFLICT	3596	3596	E -> Q (IN REF. 2).
FT	CONFLICT	3616	3616	N -> H (IN REF. 2).
FT	CONFLICT	3686	3686	V -> A (IN REF. 2).
Query Match				
Best Local Similarity 5.0%; Score 400.5; DB 1; Length 4684;				
Matches 313; Conservative 275; Mismatches 667; Indels 331; Gaps				
QY	8	EORSPEOPAGSD	---TPSELVLSHGHPAAEASGAAGDPADADPATKLPQK	-----NGQ 57
DB	1484	EERLAEQORAEERLERLAEVAALEKQRLAEHAQAQAQARE	-AKELQORIQEEVVRE	1542
QY	58	LSSVNGVAEQGDVHVQENQEGOEVEEVDVQGESE	-----DVREKDRVDEMAANS	110
DB	1543	EAAYD--AQQKRSIQEELQQLRQSSEAEIQAKARQAEAAERSLRIE	EEIRVRLQLEA	1600
QY	111	TAVEDITKQOQETSEI	-IQIIPASENNVEVMQPBASQANDVGFVKFVGFQKFTVKK	169
DB	1601	TERQRGAEQELQALRAAEAEAAQKQQAQAEERLRQVQDESQRKQAEVLA	SRVKA	1660
QY	170	D---KNEKSDTVOLLTVKKDQEGCAE	---ASVGAGDHQPSVET	---AVGESAKES 217
DB	1661	EAAAEKQKQALQALEELRLQAEAEERWLQCAEVERARQVQVALETHQ	RSABAELOSKRA	1720
QY	218	ELKOSTEKQEGTLKQEO	-SSTEIPLQAE---SDQAAEEAEKDEGEKQEKPTKSPESPS	273
DB	1721	SFAEKTAQLERSLQEEHVAQALREAEARRAQQAQAEARAEARERQ	LERWQLKANEALR	1780
QY	274	SPVNSETTSFKFFTHGWAGWKTKSPKSKEDDLTAETAKRKEQAEKVDBEKEKTEP	333	
DB	1781	LRLOAEVLQOKSL	-----AQAAEKQKEAEAREARRRQKAERQVQRELAQE	-1830
QY	334	ASSEOPEADTDQARLSADYKVELPLEDOVGDEAS	-SESKCAPL---ATEVFDEKWEA	309
DB	1831	LEKQRLAEGTAQORLAQAEQLIRLRAETEQQGQOROLLEBELARLORE	AAATAQKQKQEL	1890
QY	390	HOEVW-----AEVHVSTVKTETEEQGGG	-----EAEGVVVEGTGE	-----SLPPE- 431
DB	1891	EAEALKVRAEMEVILASKAKAAEESRSTSEKSKORLEAEAGRFRELAEEA	EARLALAEAA	1950
QY	432	-----KLAPOEVQAEAPAEELMKSRMCVSGDHTQTLDSPEEKTLPKHPEGIVSEVE	487	
DB	1951	KRQQLAEEDAARQRAE	-AERVLAEKLAAL--GEATRL-----KT-----EAEIALKE	1995
QY	488	MLSSQRIKVGQSPGLKPLFSSGLKLSGKKQKQKGGGDEBPGEYQHITHTESPSADE	547	
DB	1996	KEAENER-----LRLAEDEAFQRRRLEEBEQAAHQKADIEERLAQ	---LRKASDSSELER	2045
QY	548	QKG--ESSASSPPEPETTCTLEKGPLAPQDGEAE	-----EGTTSDEKKEKREGITPW	597
DB	2046	QKGLVEDTLQRQVQVEEILALKASFKAAGAKAELELELGRISNAEDTLRSKEQAELE	2105	
QY	598	ASPKMVTPKKRVRRPSESOK	-----EBELEKVSATLSSTDSTVSEMQDEKVTGVEE	650
DB	2106	AARQQLAAEEERRRRAEERVQKSLAAEEBAARQKAALEEVEERLAKV	-EEARSLSRER	2164
QY	651	QKPEEPK-----RRVDTSVSWEALICVGSSKKRKAASSDDGGPRTTJGGDSHR	700	
DB	2165	AEQESARQLQLAQAAQKRLQAEKAHAFAV	--QQKEQELQOTLQOEQSVLDRLRGEAAE	2222

RESULT 12				
	PGCV_HUMAN			
ID	PGCV_HUMAN	STANDARD;	PRT;	3396 AA.
AC	P13611; P20754; Q9UNW5; Q13010; Q13189; Q15123;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Versican core protein precursor (large fibroblast proteoglycan)			
DE	(Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glia)			
DE	hyaluronate-binding protein) (GAP).			
GN	CSPG2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI TaxID=9606;			

RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM V0).
 RX MEDLINE=95105188; PubMed=7528742;
 RA Nao M.F., Zimmermann D.R., Iozzo R.V.,
 RT "Characterization of the complete genomic structure of the human
 RL versican gene and functional analysis of its promoter.";
 RN J. Biol. Chem. 269:32999-33008(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM V1).
 RX TISSUE=Placenta;
 RA MEDLINE=90053982; PubMed=2583089;
 RT Zimmermann D.R., Ruoslahti E.,
 RL "Multiple domains of the large fibroblast proteoglycan, versican.";
 RN EMBO J. 8:2975-2981(1989).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM V2).
 RX TISSUE=Glial tumor;
 RA MEDLINE=95105187; PubMed=7806529;
 RT Douvres-Zimmermann M.T., Zimmermann D.R.,
 RL "A novel glycosaminoglycan attachment domain identified in two
 alternative splice variants of human versican.";
 RN J. Biol. Chem. 269:32992-32996(1994).
 RN [4]
 RP SEQUENCE OF 2711-3396 FROM N.A.
 RX TISSUE=Lung fibroblast;
 RA MEDLINE=88007514; PubMed=2820964;
 RT Krusius T., Gehlsen K.R., Ruoslahti E.,
 RL "A fibroblast chondroitin sulfate proteoglycan core protein contains
 lectin-like and growth factor-like sequences.";
 RN J. Biol. Chem. 262:13120-13125(1987).
 RN [5]
 RP SEQUENCE OF 251-347 FROM N.A.
 RX MEDLINE=93122792; PubMed=1478664;
 RA Iozzo R.V., Naso M.F., Cannizzaro L.A., Wasmuth J.J.,
 RT McPherson J.D.,
 RL "Mapping of the versican proteoglycan gene (CSPG2) to the long arm of
 human chromosome 5 (5q12-5q14).";
 RN Genomics 14:845-851(1992).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM V3).
 RX TISSUE=Brain;
 RA MEDLINE=95181355; PubMed=7876137;
 RT Zako M., Shimomura T., Ujita M., Ito K., Kinata K.,
 RL "Expression of PG-M(V3), an alternatively spliced form of PG-M
 without a chondroitin sulfate attachment in region in mouse and human
 tissues.";
 RN J. Biol. Chem. 270:3914-3918(1995).
 RN [7]
 RP SEQUENCE OF 3333-3396 FROM N.A. (ISOFORM VINT).
 RX TISSUE=Aortic smooth muscle;
 RA MEDLINE=99327053; PubMed=10397680;
 RT Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
 RL "Versican/PG-M isoforms in vascular smooth muscle cells.";
 RN Arterioscler. Thromb. Biol. 19:1630-1639(1999).
 RN [8]
 RP PARTIAL SEQUENCE.
 RX TISSUE=Brain;
 RA MEDLINE=89174663; PubMed=2466833;
 RT Perides G., Lane W.S., Andrews D., Dahl D., Bignami A.,
 RL "Isolation and partial characterization of a glial
 hyaluronate-binding protein.";
 RN J. Biol. Chem. 264:5981-5987(1989).
 RN [9]
 RP TISSUE-SPECIFICITY OF ISOFORMS.
 RX MEDLINE=96213482; PubMed=8627343;
 RA Paulus W., Baur I., Douvres-Zimmermann M.T., Zimmermann D.R.,
 RT "Differential expression of versican isoforms in brain tumors.";
 RN J. Neuropathol. Exp. Neurol. 55:528-533(1996).
 CC -i- FUNCTION: May play a role in intercellular signaling and in
 CC connecting cells with the extracellular matrix. May take part in
 CC the regulation of cell motility, growth and differentiation. Binds
 CC hyaluronic acid.

CC -i- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -i- ALTERNATIVE PRODUCTS: At least 5 isoforms; V0 (shown here), V1,
 CC V2, V3 and Vint; are produced by alternative splicing.
 CC -i- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 is expressed
 CC in normal brain, gliomas, medulloblastomas, schwannomas,
 CC neurofibromas, and meningiomas; V2 is restricted to normal brain
 CC and gliomas; V3 is found in all these tissues except
 CC medulloblastomas.
 CC -i- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
 CC -i- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -i- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -i- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -i- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -i- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
 CC -i- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL, U16306; AAA65018.1; -
 DR EMBL, X15998; CA34128.1; -
 DR EMBL, S52488; AAB24878.1; -
 DR EMBL, U26555; AAA67565.1; -
 DR EMBL, D32039; BAA06801.1; -
 DR EMBL, J02814; AAA36437.1; -
 DR EMBL, AF084545; AAB48545.1; -
 DR PIR, S06014; S06014.
 DR PIR, A29348; A29348.
 DR PIR, A30358; A30358.
 DR HSSP, P01132; IEGF.
 DR Gene: HGNC:2464; CSPG2.
 DR MIM, 118661; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001438; EGF_II.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR001304; Lectin_C.
 DR InterPro: IPR000538; Link.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR Pfam, PF00008; EGF_2.
 DR Pfam, PF00047; Ig_1.
 DR Pfam, PF00059; Lectin_C_1.
 DR Pfam, PF00084; Sush1_1.
 DR Pfam, PF00193; Xlink_2.
 DR PRINTS, PR00010; EGFBL00D.
 DR ProDom, PD000919; Link_2.
 DR SMART, SM00032; CCP_1.
 DR SMART, SM00034; CLECT_1.
 DR SMART, SM00179; EGF_CA_1.
 DR SMART, SM00001; EGF-like_1.
 DR SMART, SM00409; IG_1.
 DR SMART, SM00445; Link_2.
 DR PROSITE, PS00010; ASX_HYDROXYL, 1.
 DR PROSITE, PS00022; EGF_1, 2.
 DR PROSITE, PS01186; EGF_2, 1.
 DR PROSITE, PS01187; EGF_CA_1.
 DR PROSITE, PS01241; Link_2.
 DR PROSITE, PS00615; C-TYPE LECTIN_1, 1.
 DR PROSITE, PS50041; C-TYPE LECTIN_2, 1.
 DR Glycoprotein, Proteoglycan, Lectin, Extracellular matrix; Sush1;
 KW Signal, Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
 KW Hyaluronic acid; Alternative splicing.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 3396 VERSICAN CORE PROTEIN.
 FT DOMAIN 37 137 IG-LIKE V-TYPE DOMAIN.

FT DOMAIN 167 244 LINK 1.
 FT DOMAIN 265 346 LINK 2.
 FT DOMAIN 348 1335 GAG-ALPHA (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
 FT DOMAIN 1336 3089 GAG-BETA.
 FT DOMAIN 3089 3125 EGF-LIKE 1.
 FT DOMAIN 3125 3163 EGF-LIKE 2.
 FT DOMAIN 3163 3290 C-TYPE LECTIN.
 FT DOMAIN 3290 3353 SUSHI.
 FT DISULFID 44 130 BY SIMILARITY.
 FT DISULFID 172 243 BY SIMILARITY.
 FT DISULFID 196 217 BY SIMILARITY.
 FT DISULFID 270 345 BY SIMILARITY.
 FT DISULFID 294 315 BY SIMILARITY.
 FT DISULFID 3093 3104 BY SIMILARITY.
 FT DISULFID 3098 3113 BY SIMILARITY.
 FT DISULFID 3115 3124 BY SIMILARITY.
 FT DISULFID 3131 3142 BY SIMILARITY.
 FT DISULFID 3136 3151 BY SIMILARITY.
 FT DISULFID 3153 3162 BY SIMILARITY.
 FT DISULFID 3169 3180 BY SIMILARITY.
 FT DISULFID 3197 3289 BY SIMILARITY.
 FT DISULFID 3265 3281 BY SIMILARITY.
 FT DISULFID 3296 3339 BY SIMILARITY.
 FT DISULFID 3325 3352 BY SIMILARITY.
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 782 782 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 809 809 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1332 1332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1398 1398 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1442 1442 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1468 1468 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1663 1663 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1898 1898 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2179 2179 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2272 2272 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2280 2280 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2360 2360 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2385 2385 N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match 4.9%; Score 395; DB 1; Length 3396;
 Best Local Similarity 19.9%; Pred. No. 1.4e-06;
 Matches 431; Conservative 252; Mismatches 725; Indels 756; Gaps 102;

QY 25 VLSGHP-AEASGAGDPADAPAT-----KLPQ-----KQQLSVNG 63
 DB 786 VLLAHTLSVEAAVTSKMSWDEDTTSKPLESTEPSASSKLPPALLTTVMGNKGDIPS 845
 QY 64 VARQG--DVHVOENQGEQEEVVDVQGESEEDVREKDRVEEMANSTAVEDITKDGQ 121
 DB 846 FTEDGADEFTLPDSIQKQLEEVTDEDIAHGXFTIRFQTTSTGIAEKSTLRDST---- 901
 QY 122 EETSEIIEIQIPASNNV-EMWOPASQANDVFKVKFV-----GPKFTVKKD 170
 DB 902 --TEEKVPPITSTEGQYATMEGSALEGEVDVLSKPVSTVPOFAHTSEVEGLAFVSYSS 959
 QY 171 KNE-----KSDTVOLLTVKKDEGEAGASVAGADH--QEPSVETAVGESAKES-----E 218
 DB 960 TQPTTVVDSHSTIPUSVPIKTDGVLVPSVPEDEVLGPSQDILVIDOTRLEATISPE 1019
 QY 219 LKQSTKEQGTAKQE-----QSSTEIPLQA-----ESPOAAEEBAKDE 256
 DB 1020 TMTTKITEGTTQEEFPWKEQTAEPVAPLSSTAWTPKEAVTFLDQEGDGSAYTVSEDE 1079
 QY 257 ---GEE-----KOEKPTKSPE-----SPSSPVN 277
 DB 1080 LLTGSERVPLETTPVGKIDHSVSPGAVTEHKVKTDEVVTLTPRIGPKVSLSPGPEQK 1139
 QY 278 SET-----TSSFKFFTHGAGWRKTSFKSKED-DLETAEKKEQEAQKVEEEK 328
 DB 1140 YETEGSSTGTGFTSSLPFSFTHITQLMBEETTKTSLEIDLGSLGFLFKPKATLIEFSTI 1199

QY 329 EKTEPAS-----EQSPAEEDTQARLSADYKVELP 359
 DB 1200 KVTVPSDITTAFFSSVDRLHTTSAPKSSAITKKPLIDREPGETTDMVIGESTSHVP 1259
 QY 360 ---LEQVQ-----DLEASSEKCAPLATE-----VFDEKMAHQEVA----- 395
 DB 1260 PTTLEDIVAKETDIDREYFTTSSPPATQTPRPTTVEDEKAFQALSTPQPASTKPH 1319
 QY 396 -EVHSTVEKTEBEEQCGGEAGGVVVEGTGSLPPEKLAEPQVEQEAPEBELM--- 450
 DB 1320 PDINVIIIEVRENKTERMSD-----LSVIGHPIDSES-KEDEPCSEEDTPVDHMAEIL 1372
 QY 451 -----KSRMCVSGGDHT-----QLTDLSPPEKTLPHKEG----- 481
 DB 1373 PEPPDIIEIDLHSENEEBECANATDVTTTPSVQYINGKHLVTVTPKDPAAEARRG 1432
 QY 482 -----IVSEVM-----LSSQERIKVQS----- 500
 DB 1433 QFESVAPQNFSDSESDTHPFIATKSTAVQPNSTETTESLEVTWKPTYPETSEH 1492
 QY 501 -----PLKKLFSSGLKLGKQKGGGGDEEPEGYOIHTE----- 540
 DB 1493 FSGGEPDVPPTVPPHEEPESGTAKK--GAESVTER---DTEVGHQAHEHTEPVSLPPEE 1546
 QY 541 -SPESADEQKGESSA-----SSPEPEBETTCLEKPLEAPODGEA-----EGTTS DGE 588
 DB 1547 SSGEIAIDQESOKIAFARATEVTFGEVEKSTVYTPITVSSASAYVSEEAVALTIG- 1605
 QY 589 KREGITPM-----ASPKM---VTPKKRV-----RRPESDKKEELEKVKSAATL 630
 DB 1606 -----NPPDDLLSTKESVWEATPQVVELSGSSSIPITEGSGEAEDEDTFTMTVDL 1659
 QY 631 SSTDSF-----VSEMODEVKTVEEQKPEPKRV-----DTSVSEALI 670
 DB 1660 SQNTTDTLITLTSRIITSEFVEPATIYVSEQPSAKVPTKVFSETDTS-EWISST 1718
 QY 671 CVGSSKKARKAKSSDDEGGPRTLGGSHRAEASKKEAGTDVAVPASTQEOQO---AQG 727
 DB 1719 TVEEKRK-----EEGTTGT-----ASTFEVYSSQTSQSDQLILPPE 1755
 QY 728 SSSPEPAGSPSEGEVSTWESFKRLVTPRKKSKELEK---AEDSSVQLSTIEPSRE 784
 DB 1756 LSPNVATSSDSG---TRKFSMLTTPQSEREMTDSTPVFTTNTLENLGAQ---TTE 1808
 QY 785 ESWVSIKKIPGRKKRADKQEQATVEDSGPVEINED-----DPNVPAVPLS--- 833
 DB 1809 HSI-----HQPGVQEGTLTPRSPASVPMEOGSGEAAADPETTTVSSPSLVN 1856
 QY 834 EYNAVBREK-----MEAGNTE-----LPQLLAGVYVSEELSK- 866
 DB 1857 EY-AIOAEKEVAGTGLSPHVETTFSTEGVLSTVMDRVVAENITQTSREIVISERLGE 1915
 QY 867 -----TLVHTVS-----VAVIDGT-RAVTSVEERSPSWIS 895
 DB 1916 NYGAEIRGSTGFLPEEDFSGDFREYSTVSHPIAKEBTVMMEGSGDAFRDTQTSPTVP 1975
 QY 896 ASVTEPLEHTAGEAMP-----PVEVTE-----KDIIAETPVLTLTQTLPGKDA 939
 DB 1976 TSV--HISHISDSEGSSTWMTSAPFBEFTSSAEGSGEQLVTVSSVW-PVLPASVQK 2032
 QY 940 HDDMVTSEVD-FTSEAVTATE---TSEALRTEEV---TEASGAEETDMVSAYSOLDTSPD 993
 DB 2033 FSGTASSIIDEGLEGVGTVNEIDRRSTIILPTAEVGTAKPAVEKVEKVGSTVS--TNFPQ 2090
 QY 994 TTE-----EATPV-QEVESGVLDTEEBEERQTOAILQAVADKVEESQVPATQTVQT 1044
 DB 2091 TIEPAKLWSRQEVNPROBIESET--TSEEQIEEKSFPSPONSQATEQTIQDSQIFFT 2148
 QY 1045 GSK-----ALEK----- 1051
 DB 2149 ELKTTDYSVLTKTKYSDDKEMKEEDTSLVNMSTDPDANGLESYTTLPEATEKSHFFLA 2208

```

QY 1052 -----VEEVEDSEVLASEKEDVMPKC---PVQAGAEHLAQSEGTQAT----- 1094
Db 2209 TALVTESIPAEHVVDSPKKEESTKH-FPKGMPTIOESDDELLFSGSGSEVLP 2267
QY 1095 PESLEVPV-----TADVDHATCOVKLOOL-MEOAVAPESSET----- 1133
Db 2268 TESVVFTEVQJNNTLYPHTSQVSESTSSDKIDFNEMENVAKEVPLVSGTIFEGSSGV 2327
QY 1134 -----LTQSETNGST--PLADSDPTADGTQODETIDSQDSKATAAARQSOVTEEE-- 1180
Db 2328 TSTTLIELISDTGAGGPVTAPLPFS-TDIGHPONQTVRAWEIEQIS--RPQITTEDQDSNK 2384
QY 1181 -AATQKEPSTLPMNVPAQOEHEGEPGDVLEPPOQELTAAPPLATETVQGE--GEV 1237
Db 2385 NSTAEINETTISTDFLAR-AVG-----PEMAKEPVTSAKPSDLYEESGEGSGEV 2436
QY 1238 DMLDGEKV-----KEQEVFVHSG-----PNSOKADVTVYD-----SEWGVAGCOE 1279
Db 2437 DIVDSFHTSATTOATROESSTTFVSDGSLKHPVPSAKAVADGPFTVSWMLPLHSEON 2496
QY 1280 KESTEVQSL--SLEBGEVETDVEKEKETKPEQVSEGEQETAAPHEGTYGKPVLLDM 1337
Db 2497 KSSPPTSTLSTMTVGERTDGSPQDR-----PREFEDSTLKPRKKPTENIITIDLDK 2549
QY 1338 PSSERKALGSLGSGPSLPD--QDKACIEV-----QVQSLDTTV----- 1375
Db 2550 EDKDLILITTESITILEIPELTSKDNTIIDHTKPEVEDIIGMQTIDITEVPSRPHDSN 2609
QY 1376 -----TOTAAVEKIVTIVISENG--SPICVGAHLPAKSSAT-----GGH 1417
Db 2610 DESNDSTQVQIYEAAVNLSTETFEBSADVLAHYQTATHDESDTYEDRQGLDHMGPH 2669
QY 1418 WTLQHAEDTVPEGPESQAE-----SIPITVPAPESTLHPDQ----- 1456
Db 2670 FT-----TGIF-APSTETELDVLPATSLPT---PRKSATVPELEGKAKAKALDMF 2720
QY 1457 -----EI--SASQRESEF--EDKPDAGDADQKSTALEKVL----- 1490
Db 2721 ESSTLSDGQALDQSEIIPTLQGFERTOEYEDKKHAGSPQEPFSSGAEEALVDHTPYL 2780
QY 1491 -----KAPPELLLEKSNKILVNIQTAVDOFAR-----TEPAETHAVDSQ----- 1533
Db 2781 SIATTHLMDQSVTEVPDVEGNSNPYYTDTTLAVSTFAKLSQTPSSPLTIISGSBASGH 2840
QY 1534 TQVP 1537
Db 2841 TEIP 2844

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M34384; AAA41685.1; -.
DR PIR; A34736; A34736.
DR InterPro; IPR001664; IP.
DR Pfam; PF00008; filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neutrone.
FT DOMAIN 1
FT DOMAIN 8 314 ROD.
FT DOMAIN 315 1805 TAIL.
FT DOMAIN 8 43 COIL 1A.
FT DOMAIN 44 55 LINKER 1.
FT DOMAIN 56 151 COIL 1B.
FT DOMAIN 152 174 LINKER 12.
FT DOMAIN 175 193 COIL 2A.
FT DOMAIN 194 196 LINKER 2.
FT DOMAIN 197 314 COIL 2B.
SQ SEQUENCE 1805 AA; 198744 MW; B40BE14717E0998D CRC64;

Query Match 4.8%; Score 388; DB 1; Length 1805;
Best Local Similarity 21.3%; Pred. No. 1,38-06;
Matches 319; Conservative 233; Mismatches 543; Indels 402; Gaps 75;

QY 13 ROPAGSDT--PSELV-----LSGHGPAEASGAAGDPADDPATKLPQKNGQLSSVNGV 64
Db 440 BEPGGKQOGHPFDLDTLSATNLNPHHPTLEAKD--GESSESRSVSIPOEDSQIMEL--V 495
QY 65 AEGQDVHVOENQEGQ--EEEVNDEVDGORESEVDREKDRVEMAN--STAVEDIRKD 119
Db 496 EKEADIEVKNVENSQAQKQESGLDTEFTQ--DSQGLQKETLKAAGEBPLMSLKTIONYETA 554
QY 120 GQE-----ETSEIIEQIP--ASENNVEWVQPAESQANDVCFKVPFVGF 163
Db 555 GKENCNSTEGHGLTGLBEPEKEKQIPLKSLEKXVE-----SEKTLNGLVPVLSELLGK 608
QY 164 KFTVKKDKNEKSDTVQILITVKKDSGEGAPASVAGADHOEPPVET-----AVGSA 213
Db 609 EDTRTEDQELMSPKG--TLKRFSSLGKESQEVVAPRSKGNLLESWTAKESQHPDLPFG 665
QY 214 SKESLXKSTKQEGTQEL--KQEGSTELPLOAESDQ--AAEEKAKDGE-----EKQE 262
Db 666 AEDQMLERLVEKEDQSPRSPPEEDQACRPLQKENGEPVLEEA--BQQLIERLIEKES 723
QY 263 KEPTSPSPSSPVNSETTSFKEFTTGAWGWRKKTFSKSKEDDLTAERKKEQEAOK 322
Db 724 QESLRSPEEE--DQEGRLSQK-----ENGCPPLGEEAEDQMLRLIEKESQESLQK 772
QY 323 VDEEKEKTEPASEQOE-----PAEDTDQARLSA-----DYEKVELPLR-- 361
Db 773 SPEENORICKPLERENQKSLRYLEENQETFPVLEBSRNQRPILSLVEBEEQIVAPLEKV 832
QY 362 --DQVGDLEASSEEKCAPLATVFPDEKMEAHQEVAVHVSIVKTEBEGGGGAGBEGV 419
Db 833 SQDSLSGSL--AEENVQPLR--YLEEDDCINKSLLE-----DKTHKSLGSLERANGDS 880
QY 420 VV-----EGTGESL--PPEK-----LAPEQVQEAAPABELKSRMVCVSGGDHTOL 465
Db 881 IITPQESQVSLRPPBEDQRIYVHLEKESQEPFRSSBEEQVW--EPLSLEGENHBL 937
QY 466 TDLSPF---EKLTPKHDEGIVSEVEMLSQQRIRIVQSGPLKFLSSGGLKKLKSKQKQK 521
Db 938 SSVKEDQVNEQSLKEKESQDSKSLSD--ESQETP---GLEPK--ENASISRLAQDQD-- 989
QY 522 KKGGGDEPBGVQIHTHSPSADQKGEASASSPEEETCTLEKGLPAPD--GEAF 580

```

```
Db 990 -----ESQKLEQETQOTLRAVGNQMAVSP-----EKVDPELPKPLGNDQ 1030
Qy 581 EGTTSDEKKRGKITWASPKWVTPK-----RVRRPSESDEKBELEKVS-----ATLSST 633
Db 1031 EIARSLGKNEQSLVSLKE-KGITEVKSLETEIIEPLET-AEEDLERKSIDTQELPWS 1088
Qy 634 D-----STVSEMQDE-----VKTVGEEQKP-----EPPKRVDTSVSWEALICVSGSKKARKA 682
Db 1089 EVARETVPEDEPPGSLGSDVNDRETLTSLKESQELSLGKWNVETRVEDSQOCLQVE 1148
Qy 683 SSSDDGGGPRTL-----GDSHRAEASDKKAGTDAVPAST-----719
Db 1149 EGILOEQHQSLEVKQELPSSGNQQRWEDVVGKAVQGEAPLATTCVGTEDKAEHLRG 1208
Qy 720 ---OEQDQAQ-----GSSSPPEAGSPSEG-----EGVSWTESFKRLVTPRK 758
Db 1209 QGGEERAAAGELLQDITVGEAWSLGSEPKQRPVPAEALDNLEGGA-----LEVPAQ 1261
Qy 759 SKSKLEKAEAD--SSVEOLSTETETPESREESWVSIKKFIPGRKKRADGKQEQATVEDSGP 816
Db 1262 SPMPEVTERDEDRAGQEQDSIEVTLGLEAARTGLEL-----EQQVVGLED--P 1307
Qy 817 VEINEDPNVPAVVPVSEYNAVEREKMEACQNTLPGLLGAVTVS-EELSKTLVHTVSVSA 875
Db 1308 RHPAREEAIPPS---LGEESVKAKIAQLEGPCKEPKGALSGDILELPKTSSEALECQ 1364
Qy 876 VIDGTRAVTSVERSPSWISASVTEPLEHTAGAMPVPEVETKDIIAET-----926
Db 1365 GHEESEMGEWEAEAS-----LETSDHEGSDAPQPRPPETEDEGAQAALTAPGPKLL 1418
Qy 927 -----PVLITQTLPEGKADHDMVTSE-----VDFTSEATATETSEALTEEVTEA 972
Db 1419 EPCSPILIT-----DAHELOFQAEGIOEAGWQPEAGSEALERVEPEFGLGIPEG 1471
Qy 973 -----SGAERTDWMVSAVSOLTPSPDTEATP-----VQEVESGVLDTTEEBRQTALIQ 1023
Db 1472 LQWEGREB-----SEADLGETLPDSTPLGLYLRSPAPKWDLAGQR-----LS 1518
Qy 1024 AVADKKEE--SOVPAQTQVORTGSKALEKVEVEEDSEVLASEK-----EKDVMKGP 1075
Db 1519 PQGDACKEDWGPVPAQA-----GLSPPEEEEOGHGSDLSSEEFDLGTEASLLPGVP 1573
Qy 1076 VQEAQAEHLAQ-----GSETQATPESL--EVPEVTADVHVAT-----1112
Db 1574 KEV--ADHVQVPPVLQPCACQDGESDGFADBEESGEEGEDADEGAESGAQWWSG 1631
Qy 1113 -----COVTKLQQLMEQAVAPESSETLTDSSETNGST-----PLADSDTDGTCQDETID 1161
Db 1632 ASGGGCKVQDIAQRGDPVQESVGVSLGMDGLRGAANVPALFEMVVSQDSAPSGSEES-- 1689
Qy 1162 SQDSKATAAVROSQVTEEEAATAQKEEPSPLPNNVPAQEHBGEPGRDVLPTQBELTAA 1221
Db 1690 ---ESASLEEGEQVTDH--LDAPQEVTSVMPGVGDAFDIGGQSPNLD-----1732
Qy 1222 AVPLAKTEVG---QSGEVVDLGEKVKERQEVF-----VHSGPN 1258
Db 1733 SEQVGMKENGLEQAEQGV-VLDGD---EQELLQOQEVGALKPLVPAASPVHLGPS 1785

RESULT 14
MAP2_RAT
ID MAP2_RAT STANDARD; PRT; 1861 AA.
AC P15146;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 2 (MAP 2) (MAP2B) [Contains: MAP2C].
GN MAP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=90251471; PubMed=2339070;
RA Kinder S., Schwanke B., Schulz B., Garner C.C.;
RT "Complete cDNA sequence encoding rat high and low molecular weight
RT MAP2.";
RL Nucleic Acids Res. 18:2822-2822(1990).
RN [2]
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=91060576; PubMed=2174050;
RA Kinder S., Schulz B., Goedert M., Garner C.C.;
RT "Molecular structure of microtubule-associated protein 2b and 2c from
RT rat brain.";
RL J. Biol. Chem. 265:19679-19684(1990).
RN [3]
RP SEQUENCE OF 1-151; 1515-1694 AND 1726-1861 FROM N.A.
RX MEDLINE=90221819; PubMed=2126166;
RA Doll T., Papadrikopoulou A., Matus A.;
RT "Nucleotide and amino acid sequences of embryonic rat MAP2c.";
RL Nucleic Acids Res. 18:361-361(1990).
RN [4]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=89365159; PubMed=2770869;
RA Papadrikopoulou A., Doll T., Tucker R.P., Garner C.C., Matus A.;
RT "Embryonic MAP2 lacks the cross-linking sidearm sequences and
RT dendritic targeting signal of adult MAP2.";
RL Nature 340:650-652(1989).
RN [5]
RP SEQUENCE OF 1695-1725 FROM N.A.
RX MEDLINE=94110302; PubMed=8282767;
RA Doll T., Meichner M., Riederer B.M., Honegger P., Matus A.;
RT "An isoform of microtubule-associated protein 2 (MAP2) containing
RT four repeats of the tubulin-binding motif.";
RL J. Cell Sci. 106:633-640(1993).
CC -!- FUNCTION: THE EXACT FUNCTION OF MAP2 IS UNKNOWN BUT MAPS MAY
CC STABILIZE THE MICROTUBULES AGAINST DEPOLYMERIZATION. THEY ALSO
CC SEEM TO HAVE A STIFFENING EFFECT ON MICROTUBULES.
CC -!- ALTERNATIVE PRODUCTS: VARIOUS FORMS OF MAP2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE. MAP2C, THE LOW MOLECULAR
CC FORM OF MAP2, LACKS THE CENTRAL DOMAIN OF MAP2A/B.
CC -!- DEVELOPMENTAL STAGE: MAP2C IS EXPRESSED DURING EMBRYONIC BRAIN
CC DEVELOPMENT AND UNTIL POSTNATAL DAY 10. MAP2B IS EXPRESSED
CC THROUGHOUT BRAIN DEVELOPMENT.
CC -!- SIMILARITY: CONTAINS 3 OR 4 TAU/MAP REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X51842; CAA36135.1; -.
CC EMBL; X17682; CAA35667.1; -.
CC EMBL; X71487; CAA50588.1; -.
CC PIR; S07887; S07887.
CC PIR; S10003; S10003.
CC InterPro; IPR001084; Tubulin_Tau.
CC Pfam; PF00418; tubulin-binding; 4.
CC PROSITE; PS00229; TAU MAP; 3.
CC Microtubules; Repeat; Alternative splicing; Calmodulin-binding.
CC DOMAIN 1454 1474 CALMODULIN-BINDING (POTENTIAL).
CC REPEAT 1664 1694 TAU/MAP MOTIF.
CC REPEAT 1695 1725 TAU/MAP MOTIF.
CC REPEAT 1726 1756 TAU/MAP MOTIF.
CC REPEAT 1757 1788 TAU/MAP MOTIF.
CC VARSPLIC 152 1514 MISSING (IN ISOFORM MAP2C).
CC VARSPLIC 1695 1725 MISSING (IN ISOFORM WITH 3 TAU/MAP
CC REPEATS).
CC SEQUENCE 1861 AA; 202409 MW; 42DCF116D21EF54E CRC64;
```

Query Match 4.8%; Score 388; DB 1; Length 1861;
 Best Local Similarity 19.8%; Pred. No. 1.3e-06;
 Matches 306; Conservative 220; Mismatches 590; Indels 432; Gaps 66;

QY 7 TQORSPQAPAGSDTPELVL-----SGHGPAABASGAGADPADADPAKLLQKQKQSSVNG 63
 DB 335 TSGSQPFAVPFQSDKMSLQDTSGATSKES-----KDEBPQDKADKADVVSZA 388
 QY 64 VAEQGVHVE-----ENQEGQEEVDEVDVQORESEVDREKRDVEEMANSTAVEDIT 117
 DB 389 TTVLGDVHSPAVBGVGENISGEKGTTOQEKKETSPVQV-----PLTETEPQT 440
 QY 118 KQGEETSEIIHQIPASENNVEMVOPASQANDVGFKKVFKVGFVKFYKDKKNEKSDT 177
 DB 441 K--LEETSKV-----SIEETVAKKEESL-----KLKDKKAV 470
 QY 178 VQULT---VKQDEGEGEASVGA--GDHQBPSVETAVGSASKESELKOSTEKQEGTLKQ 232
 DB 471 IQTSTQSFQSKEDQKQEQETIEALKQDSFPILEQAVTDMAATKLTLEKXTSEPEAVSEK 530
 QY 233 EOSSTEIPLQAESDQAAEEAKDEGEKQEKPTKSP--ESPSSPVNSETTSSFKKFTN 290
 DB 531 RE-----IQGFIEDIAKSKLEGAGSATVAEVEMPFYEDKSGMSKYFETSAKEDVTR 584
 QY 291 G---WAGMRKTSFKKSKEDDLTAEKKEQAEKVDEBEKKEBPASBEOPEADTDQA 347
 DB 585 STGLGSDYVELSDRGNQAOESLDTVSPKNOQ-----DEKILLAKASQPSPPAHAGYS 637
 QY 348 RLSDAV---EKVELPLEDDQVGLASSEKCAPLATEVEDEKMAHQEVAEHVSTVEK 404
 DB 638 TLAQSTSHPSLEPRE-----PSSPQERMTTPDKVYGEKKDLHSKKNKDLTLT--- 687
 QY 405 TEEEOGGEAEAGGVVE-----GTGESLPPEKLAEPQEVPOE 442
 DB 688 --RSLGLGGRS--AIEQRSMINLPMGLDSIALGFNNGRHDLSP----- 729
 QY 443 ABPAEELMKRENCVSGDHTQTLTDLSPBEKTLPKHPEGIVSEVMLSOERIKVQGSPL 502
 DB 730 --LASDILNTSGMDEGD--DYLPTTPAVEKIPCFP--IESKEEDKTEQAKVGGQT 783
 QY 503 KTLFSSGKLKLSGKKQKQKGRGGGGBEPGEVYHHTESPESADQKSSASPEEP-- 561
 DB 784 TQVETSS-----ESPFPKEXYKNGTVAAPLPEH 813
 QY 562 -----ETTCLEKGPLLEAPQDEAEEGTTSDDEKREKREGITTPWASFKWV 604
 DB 814 LDLAGTRSLASASADAEVARRKSVSEAV---VAESST-----GLPVAADSGQPV 861
 QY 605 TPKKRVRRPESDKEELEKVSATLSS--TDSTVSEMDEKTVGE---EOKPEEPKR 658
 DB 862 KP-----DSQLEMGYCVFNKTYVPLPSPVQDSENLSGSGSFYEGTDDKVR 909
 QY 659 RVDTSVSMELICVG--SSKKRAKKAASSDDEGPRTLGGOSHRAEASKQKKEAG--TDAV 715
 DB 910 DLATDLS---LIEVKLAAGRVADDEFTAEKEASPPSSADKSGISREPDQKRAKNDLDTV 966
 QY 716 PASTQEQ--DOAGSSSPPEAGSPSE--GEGVSTWE--SFKRLVTPPKKSKLEKAED- 769
 DB 967 LEKSEHNVSKHAKSEBEVGDVLEFLGLGV--TYEQTSKAKELT-----TKETPAPAAKNG 1021
 QY 770 -SSVEQLSTEIEPSEESVWSIKK1IPGRKKRADKQEQAVEDSGPVEINEDD----- 823
 DB 1022 LSSVPEVA--EVEET-----TKADQGGIDVAAKKDDQSPLDIKVSDGQMA 1064
 QY 824 --PNVAVAVPLS--EYNAVEREKMAQNTLPLQLG-----AVYISELSK- 866
 DB 1065 SGMSVDAGKTIELKFEVDQTLTSSBAPOETDSFMGIESSHVADGAKVSETEKKEVAKP 1124
 QY 867 TLVHTVSV-----AVIDGTRAVTSVEERSPSM-----ISASVTEPLEH 904
 DB 1125 DLVHQAVDKESYESSGHESLTMSLAKPDEGKETSPETSLIQDEVALKLSVEIPLCP 1184

QY 905 TAGEMPPVEBEVTEKDI-----IAEETPVLTQTLPECKXAHNDMTSEVDTSEATATET 960
 DB 1185 PVSEADSSIDEAEVQMFQIQLPEKESTETPDIP-----ALPSTDVQPOPEAVVSEP 1236
 QY 961 SEALTEEVEENSGAEFTDMWASVQSLD--SPDTTEATPVQEVESGVLDTEEBEERQT 1018
 DB 1237 AEVREBEIEIEAEGEYDKLRLRSDTLQITDILLVPSRREF-----VETCPBEH 1284
 QY 1019 QAILQAVADKVKESQVPAOTQTVORTSKALEKVE-----EVEDESVLASEKEDVMPK 1073
 DB 1285 KGVSESVV--TLEDPFITVVQTTDEGELGSHSVFAAPVQPEEERRRPPHDELEVL-- 1340
 QY 1074 GPVQAGAEHLAQSGETQAPESLEVEPVTAADVHVATQVYIKLQQLMEQAVAPESSET 1133
 DB 1341 -MAAEQAQAP--KQGSPPADAPATPEKEVP-----PSEYKTEY 1374
 QY 1134 LTD--SETNGSTPLADSDPTADQOQDETIDSQDSKATAVAQSQVTEBEAATAQKEPST 1191
 DB 1375 YDYKDEFTTIDSIINDADSL-----WVDYQDD-----DRSLTQLEITIPKEERAE 1420
 QY 1192 LFNVPQAEHGEER---PGRD-VLEPTQQLTAAAVPLAKTEVQGEVMDLGEKVK 1246
 DB 1421 KEARRPSLEKHKRKEKPFYTGGRGRISTPERRRVAKKEPSTVARDEVRKKAV--YKKAELA 1478
 QY 1247 BEQEVFVHS-----GPNQKADAVTYDSEV 1271
 DB 1479 KESEVQAHSPSRKLLKPAIKYTRPETHLSYKRTTATSGSAQAPSAFKQAK--DKVT 1535
 QY 1272 MGVACQCEKESTEVQSLD-----EEG--EMETDVEKEKRETKPEQVSEEGOE 1318
 DB 1536 DGITSPKRSRLPRPSSILPRRGVSGDRHENSLSNLSISSARTTRSEPIRAGNSG 1595
 QY 1319 TAPAEHGTGKPVLTLDMPSSERKALGSLG--GSPSLPQDKACGI 1364
 DB 1596 TSTPT---TPGSTAITPGTPPSYSRTPTGTPSPYPRTPGHPKSGIL 1640

RESULT 15
 NFH MOUSE STANDARD; PRT, 1087 AA.
 AC P19246; O61959;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet H protein (200 kDa neurofilament protein)
 DE (Neurofilament heavy polypeptide) (NF-H).
 GN NFH OR NFH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89121513; PubMed=3220257;
 RA Julien J.-P., Cote F., Beaudet L., Sidky M., Flavell D., Grosfeld F.,
 RA Mushynski W.,
 RT "Sequence and structure of the mouse gene coding for the largest
 RT neurofilament subunit.",
 RL Gene 68:307-314 (1988).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89089138; PubMed=3145094;
 RA Sheidman P.S., Garden M.U., Lees J.F., Lazzarini R.A.,
 RA "The structure of the largest murine neurofilament protein (NF-H) as
 RT revealed by cDNA and genomic sequences.",
 RT Brain Res. 464:217-231 (1988).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Brain;
 RA Garden M.U.;
 RA Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IP PROTEINS: L, M,
 AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

CC NF-H HAS AN IMPORTANT FUNCTION IN MATURE AXONS THAT IS NOT
 CC SUBSERVED BY THE TWO SMALLER NF PROTEINS.
 CC -!- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P, NFH IS
 CC PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS
 CC THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF
 CC INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE
 CC OF AXONAL CALIBER.
 CC -!- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING
 CC OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE
 CC LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND
 CC COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.
 CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 534
 CC TO 716 AND IS SHORTER DUE TO FRAMESHIFTS.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M24496; AAA39813.1; --
 CC EMBL; M23349; AAA39813.1; JOINED.
 CC EMBL; M24494; AAA39813.1; JOINED.
 CC EMBL; M24495; AAA39813.1; JOINED.
 CC EMBL; M35131; AAA39809.1; ALT_FRAME.
 CC EMBL; M23102; CAA83229.1; --
 CC PIR; JTO3368; QFM5H.
 CC PIR; A43778; A43778.
 CC MGD; MGI:97309; NFh.
 CC InterPro; IPR001664; IF.
 CC Pfam; PF00038; filament; 1.
 CC PROSITE; PS00226; IF; 1.
 CC Intermediate filament; Coiled coil; Neurone; Phosphorylation;
 CC Repeat.
 CC FT DOMAIN 1 97 HEAD.
 CC FT DOMAIN 98 408 ROD.
 CC FT DOMAIN 409 1087 TAIL.
 CC FT DOMAIN 436 517 GLU-RICH (ACIDIC).
 CC FT DOMAIN 519 886 50 X 6 AA TANDEM REPEATS OF K-S-P-A-E-A.
 CC FT DOMAIN 887 1087 GLU/LYS-RICH.
 CC FT DOMAIN 98 129 COIL 1A.
 CC FT DOMAIN 130 141 LINKER 1.
 CC FT DOMAIN 142 239 COIL 1B.
 CC FT DOMAIN 240 261 LINKER 12.
 CC FT DOMAIN 262 283 COIL 2A.
 CC FT DOMAIN 284 287 LINKER 2.
 CC FT DOMAIN 288 408 COIL 2B.
 CC FT CONFLICT 133 133 K -> QA (IN REF. 2 AND 3).
 CC FT CONFLICT 199 199 A -> AR (IN REF. 2 AND 3).
 CC FT CONFLICT 281 281 S -> T (IN REF. 2 AND 3).
 CC FT CONFLICT 492 492 L -> G (IN REF. 2 AND 3).
 CC FT CONFLICT 551 551 P -> PREAKSP (IN REF. 3).
 CC FT CONFLICT 689 712 MISSING (IN REF. 3).
 CC FT CONFLICT 714 714 G -> A (IN REF. 3).
 CC FT CONFLICT 814 814 V -> M (IN REF. 2 AND 3).
 CC FT CONFLICT 843 843 T -> N (IN REF. 2 AND 3).
 CC SEQUENCE 1087 AA; 116612 MW; 57BAC76A38ED1CB9 CRC64;
 CC
 CC Query Match 4.8%; Score 385.5; DB 1; Length 1087;
 CC Best Local Similarity 23.8%; Pred. No. 9.3e-07;
 CC Matches 219; Conservative 139; Mismatches 374; Indels 189; Gaps 43;
 CC
 CC QY 26 LSGHGPA-AEASGAGDPADAPATKLPQNGQLSSVNGVAEGDVHVOE-----ENQEQ 80
 CC DB 240 IQCGGAQAQAQAARDALKCDVTSALRIRAQL---EGHAVQSSLSQSEWFRVRLRLS 296
 CC
 CC QY 81 EBEVDEVDGQRESDVREKDRVEEMAANSSTAVEDI--TKDQGE-ETSEIIIEQIPASENN 137
 CC DB 297 EAAKVTNDAMRQAQEITEYRR--QLQARTTEALAKSTKESLERQSRSELEDRHQADIAS 354

QY 138 VEMVQPAESQANDVGFKKVFKVGFKFTVKKDKKNEKSDTVQLLTVK-----KDE 187
 DB 355 YQDAIQQLDSELRN-----KWEMAAQLREYQD---LLNVKMALDIEIAAYRKL 400
 QY 188 GEGAAASVAGDGH-----QEPSVETAVGESAKSELKOSTEQ-----EGTLKQEQ 234
 DB 401 LGECECRIGFGSPSLTEGLPKFISITHIKVSEEMIKVVEKSEKTEVIVEGQTEIR 460
 QY 235 SETEIPQAESD-QAAEEBAKDEBEEKQE---PTKSPESPSPVNSSETTSFSSKFFTH 290
 DB 461 VTGEGVTEEDKQAQGEAGEAGEEKEEELAAATSPAAEEAASPEKETKSRVKE--- 516
 QY 291 GWAGWRKTSFPKSKSDDLLETAKEKQPAEKVDE-----EKEKTEPASEQEPADTD 345
 DB 517 -----BAKSPGEAKSPGEAKSPGEAKSPGEAKSPGEAKSPGEAKSPGEAKSP 568
 QY 346 QARLADYKVELPLEDQVDLEASSEKCAPLATEVDFDEKMEAHQEVVAHVSVTEKT 405
 DB 569 PAEAKSPAEKSPATVKSPEAKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPA 622
 QY 406 EEEQGGGGAEGGVVVGTEGSLPPEKLAEPQEVPEAE---PABELMKSRMCVSGGDH 462
 DB 623 PAEAKSPAEKSPATVKSPEAKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPA 680
 QY 463 TQTLDSLPEEKLPHKEGIVSEVEMLSQE-----RIKVGQ---SPLKLFSSSGLKKL 514
 DB 681 K-----SPAE---PKSPAEAKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSP 731
 QY 515 SGKQKQKGGGGD-----EEPGCYOH-IHTESPESADE--QKGESASSPEPEETT 564
 DB 732 AAVKSPGEAKSPGEAKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPA 791
 QY 565 CLEKGPLEAPQDGEABEGTTS---DGEKKREGITPW-----ASFKKQVTPK 607
 DB 792 SPVKEDIRKPAEAKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSPAEKSP 851
 QY 608 KVVRRPS-ESDKEELEKVKVATSLSTSTVSEMQDEVKT-VGEBOQKPEEPKRRVDTSVS 665
 DB 852 EQVKSPAKEAKSPKEEAK-----TSEKVAPKKEEVKSPVKEEVKAKEPKKVEEE-- 903
 QY 666 WEALICVSSKKRANKASSDDDEGGPRTLGGDSHRAEASDKDEAGTDAVPASTQEQDA 725
 DB 904 -----KTLPTPKTEAKSSKKD-EAPKEAPKPKVEEKKKE- 935
 QY 726 QGSSSPGAGSPGEGVSTWESFKRLV-----TPRK---KSKSLEEKAEADSSVEQLSTE 778
 DB 936 --TPTEKPKDSTAEAKKEAGEKKKAVASEETPAKLGVEEAKPKETTKTTAEADTK 993
 QY 779 I-EPSR-BESWVSIIKFIPIGRKKRADQKQEQATVEDSGPVE-----INBDDPNVPVAV 829
 DB 994 AKEPSKPTETEKPKKEEMPAAPAEKK-DTKEEK-TTESRKPEKPKMEAKVKEDDKSLS-- 1049
 QY 830 VPLSEYNAVEREKMEAQGNTE 850
 DB 1050 ---KEPSKPKTEKAEKSSSTD 1067

Search completed: December 13, 2002, 00:13:37
 Job time : 85 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2002, 00:00:30 ; Search time 94 Seconds
(without alignments)
3498.417 Million cell updates/sec

Title: US-09-902-432-4

Perfect score: 8073

Sequence: 1 MGAGSSTQRSEQAGSDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8073	100.0	1596	11 Q62766	Q62766 rattus norv
2	8030	99.5	1596	11 Q921F7	Q921F7 rattus norv
3	6512.5	80.7	1684	11 Q9WTQ5	Q9WTQ5 mus musculus
4	6035.5	74.8	1579	11 Q9MP1	Q9MP1 mus musculus
5	1805	22.4	2471	13 Q9DDN8	Q9DDN8 xenopus lae
6	642	8.0	2768	5 Q9VC00	Q9VC00 drosophila
7	635	7.9	5458	5 Q9U459	Q9U459 plasmodium
8	588	7.3	17352	5 Q95YM2	Q95YM2 procamburus
9	585.5	7.3	1829	5 Q22248	Q22248 caenorhabdi
10	574	7.1	5327	5 Q76891	Q76891 drosophila
11	564.5	7.0	2910	10 Q9FNDS	Q9FNDS arabidopsis
12	558	6.9	16215	5 Q9NFS3	Q9NFS3 drosophila
13	557.5	6.9	6815	5 Q917U4	Q917U4 drosophila
14	517	6.4	2081	10 Q91H98	Q91H98 arabidopsis
15	501.5	6.2	3484	5 P91257	P91257 caenorhabdi
16	493.5	6.1	2083	5 Q9N435	Q9N435 caenorhabdi

17	489.5	6.1	1852	3 Q9C2H4	Q9C2H4 neurospora
18	489	6.1	6632	5 Q01761	Q01761 caenorhabdi
19	489	6.1	6994	5 Q17343	Q17343 caenorhabdi
20	488	6.0	5170	5 Q17490	Q17490 caenorhabdi
21	485	6.0	2484	5 Q9U347	Q9U347 caenorhabdi
22	485	6.0	2607	5 Q23187	Q23187 caenorhabdi
23	485	6.0	3263	5 Q917U3	Q917U3 drosophila
24	485	6.0	6632	5 Q17362	Q17362 caenorhabdi
25	472	5.8	1514	5 Q85Y55	Q85Y55 drosophila
26	466.5	5.8	1871	10 Q9SRDS	Q9SRDS arabidopsis
27	463	5.7	1661	5 Q06166	Q06166 plasmodium
28	462	5.7	1786	5 Q9U0P0	Q9U0P0 plasmodium
29	456.5	5.7	1804	11 Q9Z1Q1	Q9Z1Q1 mesocricetu
30	452	5.6	3239	5 Q8T1Q2	Q8T1Q2 bombyx mori
31	442	5.5	1326	16 Q9L2C3	Q9L2C3 streptomyce
32	440.5	5.5	1812	5 Q9VZD9	Q9VZD9 drosophila
33	438	5.4	1128	5 Q26947	Q26947 trypanosoma
34	438	5.4	1616	5 Q9VSJ0	Q9VSJ0 drosophila
35	434	5.4	1510	5 Q25920	Q25920 plasmodium
36	431	5.3	3257	5 Q9V736	Q9V736 drosophila
37	430	5.3	3111	5 Q9VH10	Q9VH10 drosophila
38	429	5.3	2748	5 Q9V4P4	Q9V4P4 drosophila
39	429	5.3	3261	4 Q9Y556	Q9Y556 homo sapien
40	429	5.3	3664	4 Q9GT58	Q9GT58 homo sapien
41	425	5.3	1110	13 Q91255	Q91255 petromyzon
42	420.5	5.2	4900	5 Q9N541	Q9N541 caenorhabdi
43	419	5.2	3187	11 Q63714	Q63714 rattus norv
44	414	5.1	1422	6 Q95KU4	Q95KU4 canis famil
45	414	5.1	5085	11 Q9JKS6	Q9JKS6 rattus norv

ALIGNMENTS

RESULT 1

Q62766 PRELIMINARY; PRT; 1596 AA.
ID Q62766
AC Q62766;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SSECKS.
GN 322.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN= FISHER; TISSUE=EMBRYO;
RX MEDLINE=95257957; PubMed=7739556;
RA Lin X., Nelson P.J., Frankfort B., Tomblar E., Johnson R.,
RA Gelman I.H.;
RT "Isolation and characterization of a novel mitogenic regulatory gene,
RT 322, which is transcriptionally suppressed in cells transformed by src
RT and ras.";
RL Mol. Cell. Biol. 15:2754-2762(1995).
DR EMBL; U23146; AAA79517.1; -;
DR InterPro; IPR001573; PfamA anch.
SQ SEQUENCE 1596 AA; 172498 MW; CF78F2D161D9C7E2 CRC64;
Query Match 100.0%; Score 8073; DB 11; Length 1596;
Best Local Similarity 100.0%; Pred. No. 1.9e-314;
Matches 1596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAGSSTQRSEQAGSDTTPSELVLSHGHPAAEASGAAGDPADADPATKLPQKNQQLSS 60

Db 1 MGAGSSTQRSEQAGSDTTPSELVLSHGHPAAEASGAAGDPADADPATKLPQKNQQLSS 60

QY 61 VNGVAEQGDVHVQENQEQEEVDEVDGQRESDVREKDRVVEEAAANSTAVEDITKDG 120

Db 61 VNGVAEQGDVHVQENQEQEEVDEVDGQRESDVREKDRVVEEAAANSTAVEDITKDG 120


```
Db 241 LOAESDQAAEEBAKDEGEKEKEPTKSPSPSPVNSSETTSFFKKPFFTHGAGWRKKT 300
Qy 301 FKSKEDDLTAKRKEQAEKVDDEBEKEKTEPASEBQEPABDTQARLSADYKVELPL 360
Db 301 FKSKEDDLTAKRKEQAEKVDDEBEKEKTEPASEBQEPABDTQARLSADYKVELPL 360
Qy 361 EDQVGDLEASSEKCAPLATEVDEKMEAHQEVVAHVHSTVKTBEEOGGGEAGGVV 420
Db 361 EDQVGDLEASSEKCAPLATEVDEKMEAHQEVVAHVHSTVKTBEEOGGGEAGGVV 420
Qy 421 VEGTGESLPPEKLAEPQVQEAEPABELMKSRMCVSGGDHTQLTDLSPSEKTLPKHPE 480
Db 421 VEGTGESLPPEKLAEPQVQEAEPABELMKSRMCVSGGDHTQLTDLSPSEKTLPKHPE 480
Qy 481 GIYSEVEMLSQBRIRIKVQSGPLKKLPSGGLKKLGGKKKGKGGGDEBPBYQHHT 540
Db 481 GIYSEVEMLSQBRIRIKVQSGPLKKLPSGGLKKLGGKKKGKGGGDEBPBYQHHT 540
Qy 541 SPESADQKGESSASSPEEPETTCLEKGPLEAPODGEAEEGTTSDGKKREGITWASF 600
Db 541 SPESADQKGESSASSPEEPETTCLEKGPLEAPODGEAEEGTTSDGKKREGITWASF 600
Qy 601 KMWTPKVRVRRSESDKEBELEKVKASATLSSTDSTVSEMDEVKTVGERQKPEPKRRV 660
Db 601 KMWTPKVRVRRSESDKEBELEKVKASATLSSTDSTVSEMDEVKTVGERQKPEPKRRV 660
Qy 661 DTSVSWEALICVSSKKRKAASSDDGGPRTLGGDSHRAEASKDKAAGTDAVPASTQ 720
Db 661 DTSVSWEALICVSSKKRKAASSDDGGPRTLGGDSHRAEASKDKAAGTDAVPASTQ 720
Qy 721 EDOAQGSSSPAGSPSEGEVSTWESFKRLVTPRKSKSKLEEKAEADSSVQLSTEIE 780
Db 721 EDOAQGSSSPAGSPSEGEVSTWESFKRLVTPRKSKSKLEEKAEADSSVQLSTEIE 780
Qy 781 PSREESWVSIKKIPGRKKRKAQKQEQATVEDSGPVEINEDPNVPAVPLSEYNAVER 840
Db 781 PSREESWVSIKKIPGRKKRKAQKQEQATVEDSGPVEINEDPNVPAVPLSEYNAVER 840
Qy 841 EKMEAGNTPELPGLLGAVVSEBSKTLVHTVSAVAVIDGTRAVTSVEERSPSWISAVTE 900
Db 841 EKMEAGNTPELPGLLGAVVSEBSKTLVHTVSAVAVIDGTRAVTSVEERSPSWISAVTE 900
Qy 901 PLEHTAGEAMPVEEYTEKDIIAEETPVLTOTLPEGKAHDDMTSEVDFTSAAVTATET 960
Db 901 PLEHTAGEAMPVEEYTEKDIIAEETPVLTOTLPEGKAHDDMTSEVDFTSAAVTATET 960
Qy 961 SEALRTEEVTEASGABETTDMSAVSOLTDSPTTTEATPVQEVESGVLDTDEEERQTOA 1020
Db 961 SEALRTEEVTEASGABETTDMSAVSOLTDSPTTTEATPVQEVESGVLDTDEEERQTOA 1020
Qy 1021 ILQAVADKVKESQVPATQTVORTGSKALEKVEEVEDSEVLASEKEKDVMPKPVQEAG 1080
Db 1021 ILQAVADKVKESQVPATQTVORTGSKALEKVEEVEDSEVLASEKEKDVMPKPVQEAG 1080
Qy 1081 AEHLAQSETGQATPSLEVPETADVHVATQCVIKLQOLMQAVAPSSSETLTDSETN 1140
Db 1081 AEHLAQSETGQATPSLEVPETADVHVATQCVIKLQOLMQAVAPSSSETLTDSETN 1140
Qy 1141 GSTPLADSDTADGTQDETIDSDSKATAAVROSQVTEBAAATAQKEEPTLNNVPAQE 1200
Db 1141 GSTPLADSDTADGTQDETIDSDSKATAAVROSQVTEBAAATAQKEEPTLNNVPAQE 1200
Qy 1201 EHGEFGRDVLPTQOELTAAAVPLAKTEVGOGEVDMLDGKVKKEQEVFVHSGPNSQ 1260
Db 1201 EHGEFGRDVLPTQOELTAAAVPLAKTEVGOGEVDMLDGKVKKEQEVFVHSGPNSQ 1260
Qy 1261 KAADVTDSEVMGACQEKESTEVOSLSLEEGEMETDVEKEKRETKPQVSEEGQETA 1320
Db 1261 KAADVTDSEVMGACQEKESTEVOSLSLEEGEMETDVEKEKRETKPQVSEEGQETA 1320
Qy 1321 APEHGTGKPVLTLDMPSSERKALGSLGSPSLPDQKAGCIEVQVOSLDTVTQTAE 1380
Db 1321 APEHGTGKPVLTLDMPSSERKALGSLGSPSLPDQKAGCIEVQVOSLDTVTQTAE 1380
```

RESULT 3

Q9WTQ5

ID Q9WTQ5 PRELIMINARY; PRT; 1684 AA.

AC Q9WTQ5; DT 01-NOV-1999 (TREMELrel. 12, Created)

DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)

DE SSECKS (PKC binding protein SSECKS).

GN AKAP12 OR SSECKS.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=TESTIS;

RA Okita K., Kitamura H., Iwanaga T., Morimatsu M., Saito M.;

RT "Isolation and characterization of mouse SSECKS.";

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

RN [12]

RP SEQUENCE FROM N.A.

RA Camus A., Mesbah K., Rallu M., Babinet C., Barra J.;

RT "Gene trap insertion into murine SSECKS homologue interrupts one of

the transcription units while revealing the existence of a second

messenger exclusively express in the testis.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB020886; BAA76894.1; -.

DR EMBL; AF326228; AAK16150.1; -.

DR MGD; MGI:1932576; Akap12.

DR InterPro; IPR001573; PkinA, anch.

SQ SEQUENCE 1684 AA; 180694 MW; E569D55762FCB19E CRC64;

Query Match 80.7%; Score 6512.5; DB 11; Length 1684;

Best Local Similarity 83.9%; Pred. No. 3.8e-252;

Matches 1326; Conservative 72; Mismatches 154; Indels 29; Gaps 11;

Qy 1 MGAGSSTEQRSPEQAGSDTTPSELVSLSHGHPAAEASGAAGDPADADPATKLQKNGQLS 60

Db 1 MGAGSSTEQRSPEQAGSDTTPSELVSLSHGHPAAEASGAAGDPADADPATKLQKNGQLS 60

Qy 61 VNGVAEOGVHVHVEENQOEQEEVDVEDVQRESDREKDRVEEMAANSTAVEDITKDG 120

Db 61 VNGVAEOGVHVHVEENQOEQEEVDVEDVQRESDREKDRVEEMAANSTAVEDITKDG 120

Qy 121 QBETSEIIQIIPASNNVEMVQPAESQANDVGFKKVFKFVGFVKTKKDKNEKSDTVQL 180

Db 121 QBETSEIIQIIPASNNVEMVQPAESQANDVGFKKVFKFVGFVKTKKDKNEKSDTVQL 180

Qy 181 LTVKKDEGEASVAGDHQHPSPVETAVGESASKESELKQSTEQEGLTKQEQSSTEIP 240

Db 181 LTVKKDEGEASVAGDHQHPSPVETAVGESASKESELKQSTEQEGLTKQEQSSTEIP 240

Qy 241 LOAESDQAAEEBAKDEGEKEKEPTKSPSPSPVNSSETTSFFKKPFFTHGAGWRKKT 300

Db 240 LOAESDQAAEEBAKDEGEKEKEPTKSPSPSPVNSSETTSFFKKPFFTHGAGWRKKT 300

QY 465 LTDLSPKTLPHKEGIVSEVEMLSOERI KVGSGPLKGLPSSSGIKLGGKQKGR- 523
Db 358 LTDLSPKTLPHKEGIVSEVEMLSOERI KVGSGPLKGLPSSSGIKLGGKQKGRG 417
QY 524 GGGDEEPGEYQHHTSPESADEQKESASGPEEPETTCLEKGPLEAPQGEABEGT 583
Db 418 GGGDEEPGEYQHHTSPESADEQKESASGPEEPETTCLEKGPSEAPQGEABEGA 477
QY 584 TSDEKKEKREGITWASPKQWTPKRVRRPSESDEKBELEKVKXSATLSDSTVSEMQDE 643
Db 478 TSDEKKEKREGITWASPKQWTPKRVRRPSESDEKBELEKVKXSATLSDSTVSEMQDE 537
QY 644 VKTGEOKEPEEKRRVDTSVSEALICVGSKKRABKASSDDGCGPRTLGDSHRAE 703
Db 538 VRVGEQRESEKRRVDTSVSEALICVGSKKRABKASSDDGCGPRTLGDSHRAE 597
QY 704 ASKDKAGTDAVPASTQEQOAGSSSPAGSPSEGEVSTWESFKRLVTPRKKSLSKL 763
Db 598 ASKDKA--DALPASTQEQOAGSSSPAGSPSEGEVSTWESFKRLVTPRKKSLSKL 655
QY 764 EKAEDSSVQLSTETIPPSREESVSIKKTIPGRKKRAGKQEQAVVEDSGPVEINEDD 823
Db 656 EEAEDSGAQLASEIPPSREESVSIKKTIPGRKKRAGKQEQAAVEDSGPVEINEDD 715
QY 824 PNPVAVPLSEYNAREKME--AQNTLPLQGLGAVYSEELSKTLVHTVSVAVIDGT 880
Db 716 PDVPAVPLSEYNAREKME--AQNTLPLQGLGAVYSEELSKTLVHTVSVAVIDGT 775
QY 881 RAVTSVEERSPSMTSASVTPLEHTAGEAMPVEVTEKDI IAEETPVLTQTLPEGKDAH 940
Db 776 RAVTSVEERSPSMTSASVTPLEHTAGEAMPVEVTEKDI IAEETPVLTQTLPEGKDAH 835
QY 941 DMVTSVDFTSBAVTATETSEALRTEBEVTEASGAETTTDMVSAVSQTLSDPTTEBATP 1000
Db 836 DDVTSVDFTSBAVTATETSEALRTEBEVTEASGAETTTDMVSAVSQTLSDPTTEBATP 895
QY 1001 VOEVESGLTDEEROTQAILQAVADKVBESQVPATQTVORTGSKALKVEVEEDSE 1060
Db 896 VOEVESGLTDEEROTQAILQAVADKVBESQVPATQTVORTGSKALKVEVEEDSE 955
QY 1061 VLASEKEDVMPKGPVOEAGAEHLAQOSETGOATPESLEVPEVTDVHVATCOVIKQO 1120
Db 956 VLATEKEKVVPGPVQEAETEHLAQOSETGOATPESLEVPEVTDVHVATCOVIKQO 1015
QY 1121 LMEQAVAPESSETLTDSETNGSTPLADSDTADGTQOQDETIDSDQSKATAAVROSQVTEEE 1180
Db 1016 LMEQAVAPESSETLTDSETNGSTPLADSDTADGTQOQDETIDSDQSKATAAVROSQVTEEE 1075
QY 1181 AATAQKEEPSTLNNVPAQEGHEEPGRDVLPTQOQBELTAAAVPVLAKEVGEQGEVDWL 1240
Db 1076 AATAQKEEPSTLNNVPAQEGHEEPGRDVLPTQOQBELTAAAVPVLAKEVGEQGEVDWL 1133
QY 1241 DGEKVKE--BOEVFVHSGPNSOKAADVTYDSEVMGVAGCQEKESTEVQSLSLEGEM 1295
Db 1134 DGEKVKE--BOEVFVHSGPNSOKAADVTYDSEVMGVAGCQEKESTEVQSLSLEGEM 1193
QY 1296 ETDVEKEKRETKPEQVEEEOQETAAPHEGTYGKPVLTLDMPSSERKALGSLGSPSL 1355
Db 1194 GTDVEKEETETKPEQVEEEOQETAAPHEGTYGKPVLTLDMPSSERKALGSLGSPSL 1253
QY 1356 PDQKAGCIEVQVOSLDTTQVTAQAEVKEVETVWISSETGESPECVGAHLLPAEKSSATG 1415
Db 1254 PDQKAGCIEVQVOSLDTTQVTAQAEVKEVETVWISSETGESPECVGAHLLPAEKSSATG 1313
QY 1416 GHTLQHAEDTVPLGPSQAESPIIITPAPETLHPDLOGETISASORESEBEDKPDAG 1475
Db 1314 GHTLQHAEDTVPLGPSQAESPIIITPAPETLHPDLOGETISASORESEBEDKPDAG 1373
QY 1476 PDADGKESTAIKVLKAEPEILESKNKLVLNVIOTAVDQFARTETAPETHAYDSOTQ 1535
Db 1374 PDADGKESTAIKVLKAEPEILESKNKLVLNVIOTAVDQFARTETAPETHAYDSOTQ 1433
QY 1536 VPACRLDSREPNRCWTWKMDAKMKHPVPQPREDLQV 1571

Db 1434 VPWQADSGAQOQMLDK-----DESLQV 1457
RESULT 5
Q9DDN8 PRELIMINARY; PRT; 2471 AA.
AC Q9DDN8
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Gravin-like.
GN GL.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21095185; PubMed=11165490;
RA Klingbeil P., Frazzetto G., Bouwmeester T.;
RT "Xgravin-like (Xgl), a novel putative a-kinase anchoring protein
RT (AKAP) expressed during embryonic development in Xenopus.";
RL Mech. Dev. 100:323-326(2001).
DR EMBL; AF308810; AAG45105.1; -;
DR InterPro; IPR001573; PKinA anch.
SQ SEQUENCE 2471 AA; 267987 MW; 1DCDFD7D443D1129 CRC64;
Query Match 22.4%; Score 1805; DB 13; Length 2471;
Best Local Similarity 31.5%; Pred. No. 4.5e-64;
Matches 551; Conservative 287; Mismatches 539; Indels 370; Gaps 60;
QY 82 EEVVDVGVQRESDVREKORVEEMAANSTAVEDITKGOESETSEIIEQIPASENNVEEM 141
Db 11 EEVSTNEGPN-----MEVIPSEATPKTEIKEQGQDEAPEATN--ATESPEST 60
QY 142 VQPAESQANDVGFVKFVGVFKFTVKKDKNEKSDTVQLLTVKKDEGEAASVGADHQ 201
Db 61 EQANETQSNVEGVFKFVGVFKFTVKKDKVEKSEPVQLLTVKQKVE---VNGTENHE 116
QY 202 EPSVETAVGESASKESELKQSTKEQGTLEKQESSTTEIPLQAESDQAAEAKDEGEKQ 261
Db 117 KQN-----DVSGENVLE--TEDTKEVEKQSETDPIIQTSTREEVSGKPKSEETKL 167
QY 262 EKBPTKSPSPSPVANSSETTSFKKFTGAGWGRKTSFKKSKEDDLETAEK-----R 315
Db 168 EEEQVTPSPPTNPLVAETSSPLKKFTQGWAGLRKTSFKRKEEDHQEVEKVIISEEK 227
QY 316 KEQAEKVDDEEKE-----KTEPASSEOPBAEDTDQARL 349
Db 228 EKSDSQVKEEDGENQIIVNGETLPKEANEIESKACIDKETQPSAKELAKPEECPTDQL 287
QY 350 SADYKVV-----ELPLEDQVGDLEASSEKCAPLATVEFDEKMAHOEVVAEVH--- 398
Db 288 SVLKEDIKLENADICEVPVTESI---QTSAAEAVHVSVPTEVSEVQKKEVDTTEDNDVKVT 344
QY 399 -VSTVEKTEEEQGGGAEAGGVVVEGTGSLPPEKL--AEPOEVQEAEPABELMKSR- 453
Db 345 VSTSTATSEE-----VKEPLSEMAPDTEVICDKNDKDTTKLAQTATLLESTE 393
QY 454 -EMCVSGGDHQTLDLSPBEKTLPHKEGIVSEVEMLSOERIKVQSGPLKGLPSSSGLK 512
Db 394 VEICL-----EPIMTETELSSQEKAKLQGSPLKPLFGSGSLR 431
QY 513 KLSGKQKQKGRGGGDEEPGEYQHHTSPESAD--EQKGE--SSASSPEPEETTCLEK 569
Db 432 KLSGKKNKGGK---EDDTKAENT--TEVPVSSDIPESDGDNTSPSSPEDSAETSPTK 485
QY 570 PLEAPQDGEAE--EGTTSDEKKEKREGITPWASFKQWTPKRVRRPSESDEKEELEKVKSA 628
Db 486 AEDATQAVETEGDGTSDGERKEGITPWASFKKLVTPKRVRRPSESDEKEELEKVKSA 545

QY 629 TLSTSDTVS-EMQDEVTGEEOK-----PEPKRRVDTSVSWELICVSSSKTARKRS 683
D6 546 TWSSTDSAGSVNOEAEKNGEOKLEKSTENKKVDSVSMELICVSSAKKARKRS 605
QY 664 SSDDGGRITGGDHRAEAKDGAGTDAVPASTOEDODQOGSSSPBPASPSGEGV 743
D6 606 DSDDE-TQKNIDENKKEIEVVKSKHESEDPVSSQKEQVNDIPSPQSTSPTEGDOG 664
QY 744 STWSEFKRLVTPRKKSLSKEKAD-----SSVEQLSTEIEPSREBSVWSIKK1PGRKK 799
D6 665 STWQSFKLVIPIRRKSRTAEKTEETPAVNTQOPTSDGTGKEGWSFKKL1PGRKK 724
QY 800 KRADCKQOQATVEDSGP-----VEINEDPNVAVVPLSTYNAVEEKMAQNTLIPOLL 855
D6 725 KKSDDKQEPATISDTGKSTEGCEAMEDDADVAVVPLSFDAAEQEKRAQCKEPTLNDV 784
QY 856 G--AYVSELSKTLVHTVSVAVIDTRAVTVEERSPSMISASTPELEHTAGAMFVPV 913
D6 785 SKDEKTPVNSGELIHATTVVGBERAVTSLBDAPSMISANTVETVEQANLTPRVT 844
QY 914 BEVTEKDI1AERTPV---LTQTLPGKADHDMVTSEVDFTEAVATETSEALTEEVT 970
D6 845 KERISEITVEBAF1FGEVSGVMSG-----NTINEVELTSEALTA-----LEBA1 891
QY 971 PASGAETTDVMSVANSQLTDSPTTBEATPVQEVBSGVLDTEEBEROTALLQAVADKVK 1030
D6 892 EYSCAEETTEMISAASOLDSTFTPAEEVTP1PEDEBGOSQSLDMPKQMDPVLE-ATELT 950
QY 1031 FES-----QVPRTQVQRTGSKALEKEVE-----EDSEVLASEKEKQMP-- 1072
D6 961 EOSTISLDTKIQNESISTSAEHDTKVKDKEANICISCKOBEESTIPSEDEK1ECTPVL 1010
QY 1073 -----KGPVQAGAEHLAQSGETGQATPESLEVEPTADVHV--ATCQVYK1QO 1120
D6 1011 AERHDPELSKSVESNGA-----DVSVS-----VEH1IDAV1ELVKERY 1050
QY 1121 L-MEQ--VAPESSETLTD-----ETN-GSTPLADST--ADTQODETIDSD 1164
D6 1051 LSSDDAD1YAVQDTEKHPDAVLEVNREBETGVSVEQDSSTI1ILDSSQADK1VDTAQ 1110
QY 1165 SKATAVAGOSQTEBEAATAQKEP-----STLPNNVPAQEEHGE--PG 1207
D6 1111 KQSEITVLSEKKEVBCA1MSDHTK1ASVHPKQCFEVTSVSTEDD1PVEKQVDSN1SA 1170
QY 1208 RDVLEPTQOELTAAVAVPVLAKTEVQGEVDM1DGEKVEQ----- 1249
D6 1171 VKLVVAAEQVTEKSIPTLEQKELSEBSSSTISVSTVEHQVKEALISDQDALPVK 1230
QY 1250 EYVHSGPNSQKADVYDYSEV-----MGVAGCOEKESTEVQISLS 1289
D6 1231 EQVLNDS1SAVVLVAAEQVTEKSIPTLEQKELSEBSSSISVSTVEHQVKEALIS 1290
QY 1290 LEEGEMDVVEKREKTRPEQVSEGEQETAPRHHGTVGKVL1LDMSSRGKALGSL 1349
D6 1291 DQETTERHPHASEVETK-EGAAEQVEKSLIYEVESAQVEBCSLFATQK-- 1345
QY 1350 GSGPSLPQD--KACGIEVQOSLDTVTQTAENVKY-----ETV1SETGSGPECV 1401
D6 1346 -ESVSI1LEELARKSDPVEPLEQS-GITLATYEDTTERK1IPADKETATASVAG1GSAKA 1403
QY 1402 GAHL1PAEK-----SSATGHWI---LOHADTVPLGPESQASIP1ITVPABST1H 1451
D6 1404 LNALFENKHS1LEDVSEVKGQVSCSAVMPDEQAVEAQSTEKQ--VAKSP1PA1A- 1460
QY 1452 PDLQGE1ASQERSEER-----DKPDAGDADAG 1480
D6 1461 --LTE1IAS1VEAEVKEKNAYPSEFQ1S1PSSL1PDQKEVTT1AEKDT1E1LDE 1518
QY 1481 KEST1A1E--KVLKAPPE-----1LELESKSNK1VLNV1QTAVDQFARTETAPETHAYDS 1532
D6 1519 TEKQAMESVSV1SGQEVNPNFNSV1VEEVS1KAF1AGVQ1TJESW-----1T1ASEE 1569
QY 1533 QTVQFACRLDSREPNRCWTMKDAKMKH---PVQPREDLQV1TVLEMAQPRKCLPRL 1588

D6 1570 QTR-----ECTTADVQAVECNA1TVDESEERPSAT1SHSAE---TVSV 1613
QY 1589 QUKAPVS 1595
D6 1614 SLECPIT 1620
RESULT 6
Q9VC00
AC Q9VC00 PRELIMINARY; PRT; 2768 AA.
ID Q9VC00
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE CG13648 protein.
GN CG13648.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle K., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Ashbyant A., An H.-J., Andrews-Piankocch C., Baldwin D.,
RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bandhart D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Broksrein P., Brottier P.,
RA Butris K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dublin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong P., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitel M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Kechum K.A.,
RA Jimeli B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazzolo M., Platten G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skipski W.P., Smith T.,
RA Splyer E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein D.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.-N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AE003750; AAP56376.1; -
DR FlyBase; FBgn0039257; CG13648.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00093; VWC_1.
DR SMART; SMO0214; VWC_4.
DR PROSITE; PS01208; VWF_C_1.
SQ SEQUENCE 2768 AA; 294032 MW; CA929A21774E4684 CRC64;

Query Match

8.0%; Score 642; DB 5; Length 2768;

Db 1480 EASGETSEDNEIDAGASSTPVPSADEDKTPSTKTEVADKFTTVAPLAGDEESNLP 1539

QY	833	QUGAVVSEBLSKIDVHIVSVAVLDGKAVLISVBER-----	1597
Db	1540	KLQDQIF---EEAPAVATTAAPSKDQGEQKPVVEBEKPIEDGQKPIEDTSTPTSSENEI	1597
QY	890	-----SPSMISASVTPLEHTAG-----EAMPVBEVTEKDIILAEETP---VLQTQLPEG	936
Db	1598	EPESDRATTIAPSKEEPSPSTGATKDEPAPSTDAPESD-ESKETPESEVPTTVPAPAG	1656
QY	937	KDAHDDMVTSEVDFTSEAVTATETSEALRTEEVTE-----ASGAEE---TTDMVSA---	984
Db	1657	EKIPTSSITPDEEPTATSAVAKPEDVEKETSTETPTDAPASSEEDSNSDQIIPSEVP	1716
QY	985	-----VSQLTSDPDTTEATPVQEV-----ESGVLDTSE---1013	1013
Db	1717	EKKPETPAQTPBEGDVTGATAAPTSTDSVPPVQRLPEEVLAEIPOPSTETGIKQODETTA	1776
QY	1014	-----EEROTQILQAVADKKVEESQVPATQTQVORTGSKALEKVEEVEED	1058
Db	1777	APSIDRKEPVTVIIDEATTVAPISEKDEKPTEE-----EKPVQKPTGEEP	1823
QY	1059	SEVLASEK---EKDMPKGPVQ---EAGAELHAQSETGQATPESLEVPEVTADVHDVATCOV	1115
Db	1824	SEBEEKKPTIEQDVSTEGPVS TEASEAGSTSESSEVKFSTEG-EVAEKPEDKQPSSTAQA	1882
QY	1116	IKLQQLMEQAVAPESSETL---TDSETNGSTPLADSDTADGTQODETIDS-----QD	1167
Db	1883	-----PVETIPEITSTELPAQDGKPTSEAPVDSDEDTAPSDEKIPSVSGEEVGEPEV	1935
QY	1168	TAAVROSQVTEBEEAATAQKEEPTLNNVPAQEEHGEHGPGRDVLLEPTQOELTAAAVPVL	1227
Db	1936	TTASPOA-AEEDELTKTPASEPSS-TDKVPETEQKPE-----DEYKADETPESVTQVS	1987
QY	1228	KTRVGQEGEVDMLDGSK-----VKSEQEVFVHSGPNSKOAAADVTVYDSEWGVAGCQE	1279
Db	1988	DVATSTAPVAGDIEKRDQATTAGPESEEEI---KPTIAPAAEIPQS-----E	2034
QY	1280	KESTEVQSLSLEGEOMETDVEKTR---ETRPEQVSEGEQETAAPHEG-----	1326
Db	2035	KEPVDEQ--EVESGTKATPAESDGGQPIDEIAPATSGPIDEASTAAPTKEESTTVASAASP	2092
QY	1327	-----TYGKPYVL---TIDMPSRBERGKALGSLGPSLPDQKAGCTEVQVQSILDT	1373
Db	2093	AVHDDDEIKOVTTITQPVADKEVAQAQDETKYSIDVSTDSPTAQDDEKQDKEAPV--APT	2150
QY	1374	TV-----TOTABAEKV-IETV-----VJSTGESPECVGAHLLPAEKSSATGG	1416
Db	2151	TVSSPTADSAADSSTPTVEVPSPVIDTKPMDIMSQTIAPHTADGAASTSTEDQAPV	2210
QY	1417	HWTLOHAEDTVLGPESQ-----AESTPIIIVTAPPE-----STLHPDLQGEIS	1459
Db	2211	TVSPQDAEKT-PVSPAPQSDKTPSEAPQDADEIPATATPLDUNKIIPATVAPQTDGVP	2269
QY	1460	ASQORSESEEDKPDAGP-DADGKESTAI-----EKV-LKAEPEIILELSKSNKIVLNIQT	1513
Db	2270	ATAAPLDEDKIQITTAAPLDEEKIPSTAAPLDDDEKIPAPSPVVFDEPSESSEK-----	2322
QY	1514	AVDOFARTE-TAPETHAYDSQTVQACRLDSEGNRCWTKMKDAMKHVPQPREDLQVL	1572
Db	2323	AVSEYDGEESTEPVPHVDVETSTDEPT-----SDAKLKPPTSAPATPSE-S	2366
QY	1573	TVLEAWAQPRKCLPRLOLQKAP	1593
Db	2367	PATEAIVPETAAPELEKEVP	2387

RESULT 7

Q9U459

PRELIMINARY;

PRT; 5458 AA.

AC

Q9U459; Q9GFG4;

DT

01-MAY-2000 (TREMBLrel. 13, Created)

DT

01-JUN-2001 (TREMBLrel. 17, Last sequence update)

```
RESULT 7
Q9U459
ID Q9U459 PRELIMINARY; PRT; 5458 AA.
AC Q9U459; Q9GFG4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
```

DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
 DE Erythrocyte membrane-associated giant protein antigen 332.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NC NCBI_TaxID=5833;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=FCCL/HH;
 RA Shan Z.X., Yu X.B., Li X.R., Ma C.L., Fang J.M., Lu J.H., X J.J.
 RT "Molecular cloning and structure analysis of the Plasmodium falciparum
 erythrocyte membrane-associated giant protein Ag332 (Pf332) gene."
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF202180; AAF15293.3;
 DR InterPro; IPR001313; Ptm110/Puf.
 SQ SEQUENCE 5458 AA; 615269 MW; 4F6A37D92BD0172 CRC64;
 Query March 7.9%; Score 635; DB 5; Length 5458;
 Best local similarity 23.4%; Pred. No. 5.6e-17;
 Matches 411; Conservative 292; Mismatches 688; Indels 366; Gaps 79;
 QY 4 GSSTQRSPEQPA--GSDTPSELVLGHPAAASGAADPADADPATKLPQKQQLSSV 61
 DB 977 GSVTQLVBEKEGLINDEKELIT--EMSEIKSVNEIETDLSTELIKQGLATY 1034
 QY 62 NGVAEQ--GDVHVOEN--OEGOE--EVDVEDGQRESEDRKDRVEKMAAST 111
 DB 1035 EFVEEKSILTDLKEESVTEKEVGETELSTQEVNDEKVSVEEVEEKSVEEVLKES 1094
 QY 112 AVEDITKQO-----EETSEIIQIPASE----- 135
 DB 1095 AEEVEERGSCTEIVEKSGDTEIVEEGSATEIVEEGSATEVEEVSVEEMLKEBG 1154
 QY 136 NNEVEVPAESQANDVGFKKVFKFVFKTYKDKNEKSDTVOLLTVKD-----E 187
 DB 1155 SATTEVEEGLSSDNVQSKGIENVGELIYSVTKADESMN--EKIPLKSFVDESEFK 1212
 QY 188 GEGASVAGAGD-----HOEPSEYAV-----GESAKSESL--KQ 221
 DB 1213 GGGPDNVNSEDVNEEDLINEHTPLEETKIEELPTEYITADIHTEGETETKXNLYEKI 1272
 QY 222 STKEKEGTLKQOSSTEIPLQAS--DQAEBAKDEGEEKO--EKEPKSPESPSS 274
 DB 1273 NEEVEKAFQOEKITEENTIVERESTVEDIVQEPSLAQAEVQESDTNELTEKLANEKI 1332
 QY 275 PNVSTTSFFKFFTHGMAGWRKTSFKSKEDDLTAKRKQEA--EKVDEBEK 330
 DB 1333 PEVSTENVVK-----EGDTEEVLEDESITEIVEEVSACBEIVEEESSS 1382
 QY 331 TEPAEEQEPADTQARLSADYEKV--ELPLEDQVGLNASEKCAPLAT--EVPE 385
 DB 1383 ERIVEEESSESEIVEEESSESEIVEEESSESEIIEVSTESTVLE 1442
 QY 386 KMEAHQEVAAHVSTVKT-----EEGGGGG--ABGGVAVVEGSELPE 431
 DB 1443 EGSVTEELVEE--EVSTTEVKDIGVSEVLEEBGTEPEVQOQSVLEIETESSTE 1501
 QY 442 KLAEP-----QVPOEABE--AEELMKSRMCVSGDHTQLTDLSPREKLPKHPGIVSE 485
 DB 1502 KPLEDVGSNVEIIVQEGVADQEIHEE--VSTTEKHVDVNSTTEIIEVK--VQSVSE 1556
 QY 486 ---VEMLSQERIIVQGSPLKTLFSSSGI-----KKUSGKKQKGRGG 526
 DB 1557 EIIVEEVSASEEIVEEGSVTEEEVEBEKLINEGATESVTEIIVQKESVDAEVLQOEGS 1616
 QY 527 GDEEFOEYHHTTE--SPESADQKESASSSEPEPEFTCLEKGLAPQDQAEAEGETT 584
 DB 1617 MNEELLEKSIIVEIIVGPEGSVTEELVDVGSFAEEVKEELVTE--EAVQ--YEGSV 1669
 QY 585 SDGEKKREGIT-----PMASFKMVTPPKRVRRPSSD--KEEELKVSAT--LSSTDST 636
 DB 1670 TEELKEEESITENEALAESAPAEIIEK--GPNDELVEKEGLDTEIIVNEVSTDEV 1725
 QY 637 VSE---MODEVKTGEEQKPEEKKRRVDTSVSWEALICVGS--KKRARKASSSD--D 687

DB 1766 IEEKLVNEQI--VGEESVTEKEVEVERSAT--EDLVEEGASVTEKSVHGGSTTEQLD 1782
 QY 688 EGGPTLGGDSHRAEASKDKAEAGTDAV--PASTOEDQAGSSSPER-----AGSPS 738
 DB 1783 ESVAEIIVEEVSVDKIIIEEVSVDVEVEEGSVTEIIEEESVPEIIEBELSGSEE 1842
 QY 739 -----EGEGSVTSESPFKLVTPEKRSKSKLEKASDSVEQLSTIEPGRBSW 787
 DB 1843 VLEDEWYTDAPMGQEGSVTEIEIEEIV--DDEGSTTEIIVDGSNKEIIVEEPRVEEV 1899
 QY 788 VSIKKFIPGRKKKADQKQEQATVEDSGPV--EINEDDNPVAVVPLSYNVRERKMAQ 846
 DB 1900 LKRGFTVLEIIEEGSVTEQ--VEDIKTVSEKSESSAIEEVEKVEEBSISEKIYEKE 1956
 QY 847 GNTLEPQLLAGVAVVSEELSKTLVHTVSAVVDGTRAVSVSEERSPSWISASVTEPLEHTA 906
 DB 1957 ES-----VTEELVRQEESTTEKIVKDVSPTEDFVQ-----TDSVTEKVIQOE 1999
 QY 907 GPAMPVEEVTETKDIIEETPVLTQTLPEGKADHDMDVTSYVDFTSEAVTATSEALRT 966
 DB 2000 GSNTEVADVEEKESASDE-----HEQEDVSVNAQVTECKKSVTEIYDEVERT 2048
 QY 967 EEVTEASGAEETDMVSAVSQLTDSPTTEATPVQEVESGLDTEEBERQTOALQAVA 1026
 DB 2049 EIIVEENG--VTBGSVDETGSVTE--EIIIEATVTEEVEDGSVTEEVEDGSVIQEYVE 2104
 QY 1027 D-KVKEESQVPAQTQVQTGSKALEKV-----EEVEDESEVLASEKEDVMPKPVQOE 1078
 DB 2105 DGSVTEE-----IVQNGSVTEIIEEBSGVNEEVEEVSSEVNDTEVTE--EVEE 2156
 QY 1079 AGA--EHLAQSEGTQATPESLE-----VPEVTADVHVATTCQVIKLOQLMEQAVAPSS 1131
 DB 2157 EGSVVEEIVE--EEGVSVEEIVEEBSGVVEEIVEEBSGVVEEIVEEBSGVVEEIVEEBS 2214
 QY 1132 ETLTDETSNGSPLADSDTADTQODETIDSDQSKATAVAQ-----SQVTEEAATQ-- 1185
 DB 2215 VVEEIVEEBSGV--VEEIVEEBSGVVEEIVEEBSGVVEEIVEEBSGVVEEIVEEBSGVVEE 2273
 QY 1186 --KEEPTLPPNNVPAQ-----EEHGEPPGRDYLEPTQOELTAAVPLAKTE-----VG 1232
 DB 2274 IVEEBSGVVEEIVEEBSGVVEEIVEEBSGVVEEIVEE--GSVVEIIVEEBSGVVEEIVEE 2331
 QY 1233 QGEVDMLDGEKVEKEQEV--FVHSGPNSOKAAD--VTYSEVWGVAQCOEKESTEVOS 1287
 DB 2332 BEGSV-----SEVDETLELVNDEIVEQAPFTEVEEVOVSVDDEIIEEDSVAE----- 2378
 QY 1288 LSLSEGMENTVEKEKRETKPQVSEEGEOETAPAE--HSGTVGKPLVTLDMPSERGX 1344
 DB 2379 -AVEESESITTESQOEBETEKGFVLEKVEETGAVTEEIVQDLITEE1--LESESUNGE 2435
 QY 1345 ALGSLGSPSLPDD-----KAGCIEVOVQSLDTVTQTAEAVKVIETVVISETGE 1396
 DB 2436 IINKESDAEELIETFLNEBVQAGSSEBEIVEEBSGVTEKEVESVTEELV--DEGSV 2494
 QY 1397 SPECVGAHLLEPAKESAT-----GGHWTLOHAEPTVLPQPSQAESPIIYTPAPESTL 1450
 DB 2495 TEELV-----DEGSVTEEVEEGGSAIQEIVEE-----ESATEEI-----IRDETIV 2536
 QY 1451 HEDLOGETISASQRESEEDKADAP--DADGKESATIEKVLKAEPEIIELESKSKIYVN 1509
 DB 2537 EEVLKEGSA-----TEELVQDGSNTDFVGKQSVIEEV--EEIISTTEK-----LK 2584
 QY 1510 VIOTAVDOFARTETAPE 1526
 DB 2585 EASALIEEVEEBSIRE 2601
 RESULT 8
 ID 095YM2 PRELIMINARY; PRT, 17352 AA.
 AC 095YM2;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)

DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE I-connectin.
 GN I-CON.
 OS Procambarus clarkii (Red swamp crayfish).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
 OC Astacodea; Cambaridae; Procambarus.
 OX NCB_TaxID=6728;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21423462; PubMed=11532946;
 RA Fukuzawa A., Shimamura J., Takemori S., Kanzawa N., Yamaguchi M.,
 RA Sun P., Maruyama K., Kimura S.;
 RT "Invertebrate connectin spans as much as 3.5 micrometer in the giant
 RT sarcomeres of crayfish claw muscle.";
 RL EMBO J. 20:4826-4835(2001).
 DR EMBL; AB055861; BAB64297.1; -.
 DR InterPro; IPR000577; FGGY_kin.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR01452; SH3.
 DR Pfam; PF00047; fn3; 5.
 DR Pfam; PF00018; SH3; 1.
 DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_3.
 DR PROSITE; PS00002; SH3; 1.
 DR PROSITE; PS00002; SH3; 1.
 SQ SEQUENCE 17352 AA; 1962348 MW; 4BA157BEC042B42D CRC64;
 Query Match 7.3%; Score 588; DB 5; Length 17352;
 Best Local Similarity 22.3%; Pred. No. 1.6e-14;
 Matches 381; Conservative 251; Mismatches 637; Indels 438; Gaps 76;
 QY 41 DPADADPATK-----LPKNGKOLSSVN-----GVAEQGDVHVOE 74
 DB 14336 DLADRPKTEIAKSLLPQDSIAVQSVKAPVDIKELPKSETAKPVVSTQESLSVE 14395
 QY 75 ENQEQGEVV-DEDVGQRESDVREKDRVEEMAANS-----TAVE 114
 DB 14396 VQVEGQTEMTPKQISETARSVTKDSVQVLEVSSELAPDDIPETRLIEDKAKVTVE 14455
 QY 115 DITKQGEETSEITIEQIPASENNVEWQPAESQ-ANDVGFKVFKFV--GKFTVKDK 171
 DB 14456 NLRADQEDV--IVLGPTEKIIIVSESDDELVKELPEVKEEIVEEQKLPKPTKK 14513
 QY 172 NKSQDTVQLLTVKDB--GEGAEASVGAGDHQPSVETAVGESAKSELKQSTEQEGT 229
 DB 14514 KVKEETEEVLIKPEKEEVEEIVKPKDIEIVTEVMIKKPKKTEVTEVVEEVT 14573
 QY 230 LKQSGSTIPL---QAESDQAABEAKDEGEKQ-----EKPTKSPESPSPV 276
 DB 14574 IKKPEVIEEVTVKPEKIEEITEEVIKKPEKVVVEITEEVIKKPKKPEPKIEEV 14633
 QY 277 NSETSSFKFFTHGWAGRKTSFKSKXEDDLTAKEKEQAEKVDEE-----EKK 330
 DB 14634 TEVT-----LKKPEEKIAEVEETEEIIVKKAEEKVEEVTVEQIKRPPKK 14679
 QY 331 TEPASEE-----QSPADTDQARLSADYKVELPLEDQVGDLEASSSEKCAPLATEVF 383
 DB 14680 PEPKIEEVEEVIKKPEEVEEVIKKPEVVEEVEEVEV--TIKKPEKVVVEEVEE 14737
 QY 384 DKMEAH-----QBVAAHVSVTVKTEEBQGGGGEAGGVVVBGTGE---SLPPEKL 433
 DB 14738 QIKKPKPTPEPKIEEVEEVEE---TLKPEEK-----VVEEVTVEVVIKKPKKK 14784
 QY 434 ASP--QEVQEA---EPAEELMK--SREMCVSGDHTQLTDLSPREKTIIPKPEGISVSV 486
 DB 14785 PPKPEEVEEVTIKPEEKVIEEVEEVIKKPEEKVHEEVS--EVTIKKPEEKVIEE 14843
 QY 487 EMLSSQERIKVOGSLPKLFLSSSGLKLGKQKGRGGGDEEPOEQYQHINTESPE-SA 545
 DB 14844 -----TEEVKIKRPPKK-----PEPKI 14861

QY 546 DEQKGESSASSPEEP--ETTCLERKPLAPQDGEAEBOETTSDGKKREGITPWASFMM 603
 DB 14862 EEVTEEVLTKPEEKVIEEVT--EEVIRKPEKVVEE-----14897
 QY 604 VTPKKVRPSPSDKEEEL--KVKSATLSSTDSTVSEMDQDEKTVTGEQKPEEKRRVD 661
 DB 14898 VSEETIKKPEEKVVEEVEEVIKKPKKPEPKIEEVEEVT--TI---KKPEE---KID 14950
 QY 662 TSVSWEALICVGSKKRKARKASSDDDEGPRTLGGDSHRAEASAKDKEAGTDAVPASTOE 721
 DB 14951 EEITEVEQIKRPPKKPEPKIEEVEEVIK-----KPEKVVEEITEEVIKKPE 15001
 QY 722 QQAQSSSPPEAGSPSEGEVSTWESFKRLVTPRKKSUKLEEKAEKSSVQLSTEISBP 781
 DB 15002 EKVPEEVEEVTIKKPEEKVVEEVEEVIKKMLKKKPEPKIEEVEEVTIKKPEKIDE 15061
 QY 782 SREESWVSIKKIPGRRKKRAGDKQEQATVEDSGPVEINEDDPNVAVPLSEYNAVER- 840
 DB 15062 EITEE-VQIKR---KPKKKPEPKIEEVT-----EVTLRKPKBEAIEITEEVIKKP 15110
 QY 841 -EKMEAQGNTELPQLLGAVVSEELSKTLVHTVSAVIDGTRAVTSVEERSPSWISASVT 899
 DB 15111 EEKIE-----EVSEEVIKKPEKVVEEVEEVIKKMKPKKPEKPEEVEEVE- 15162
 QY 900 EPLEHTAGE-----AMPVVEEVT-----KDIIEETPVLTQTLPEGKAHDDM 943
 DB 15163 KPEEEVAEEVQIRKPKKSVFKTEISEITILQPKPEKVEEVEEVTIKRPEK----- 15216
 QY 944 VTSEVDFTSEAVTATSETSEALRTEEVEASGAEETDMVSAVSQLTDSPTTEEATPVE 1003
 DB 15217 IVEEV---SEETIKKPEKVVVEEVT-----EVOIKMKPKMKPKYEEVEEVEE 15267
 QY 1004 VBSGVLDTSEERQQAIIQAVADVKEE---SOVPATQTVQRTGSKALEKVEEVEE 1060
 DB 15268 PE-----EVEEITEEVPFLKPKKPEKPEEVEEVEEVTIMR---KREDKVEEVEE 15318
 QY 1061 VLAESKEDVMPKPVQEAAGAEHLAQSSTGOATPESLEVPEVTDVHDVATCQVILQ 1120
 DB 15319 IKMKPEK---POEKFDDVSEETITSKPETKQVTE---EITEDVQ-----IKLK- 15362
 QY 1121 LMEQAVAPESSETLTDSETNGTPLADSDTADGTOOD--ETID-SQDSKATAVROSQVT 1177
 DB 15363 -----PKKPDITEEVEEVTIIRKEEPEEEDIESETVOIKRKKRKITTEEEVT 15415
 QY 1178 EBEAATAQKEEPTLNNVPAQEEHGEPEGDRVLEPTQOELTAAAVPVLAKEVGEVEV 1237
 DB 15416 E---VTIQKPKP-----EPEVEEVEEFLMKRPPKPKPEEPOEVEEVTI 15459
 QY 1238 DWLDGE-KVKEE--QEVFVHSGPNSQKAADVTYDSEVMGVAGCQEKSETEVQSLSEEGE 1294
 DB 15460 QKESPEPTVTEEVEEVSFRVKPKKPIVTEEISEEFTLRKPKPE-----EEGP 15510
 QY 1295 METDVEKE-KRETGP-----EOVSEGEQETAPE---HEGTYGKPVLTLDMPSSERGA 1345
 DB 15511 YEVSVEEVIKKPKPPIRETOEISEEVEEMLTEPEDFTHEETEDVKIKMKRPPKK- 15566
 QY 1346 LGSLGSPSLPDQDRAGCIEVQVQSLDITVTOTAEAVEKVIETVVISETG--ESPVCVGAH 1404
 DB 15567 -----YSVTB-----ETIDSKWIOESDEEPEEVIYE 15594
 QY 1405 L-LPAEKSSATGGHWTLOHADVTPLGPESQAESIPITVPAPESTLHPDLQGEISASQ 1462
 DB 15595 LEDVVVQKKIIKDG--TEIAQEFVQLRKKPKKSVTV-----15629
 QY 1463 RERSSE-----EDKPDAGP-DADGKESTAI-----EVLKAEPEILESKNKIVL 1508
 DB 15630 RDEGETVVIKSIITDETDAQPKHEDIEESFQLRPPKKKYKVEEVEEVEGVKIRREPT 15689
 QY 1509 NVIOAVDOQFARTET-APETHAYDSQTVQVACRLDSRPNRCWTWKDKAKMHPVQPRE 1567
 DB 15690 FEVRDEEVSFRKWTNEFE-----EPEEEVEE-RDVEGKFTFDLPK 15731
 QY 1568 DIQVLTVLAEWAQPRKCLPRLQLKAPV 1594

Db 15732 KIQKRTTEIEIYQVKV-LPPLPKQPPV 15757

RESULT 9

Q22248 PRELIMINARY; PRT; 1829 AA.

AC Q22248;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE T06D8.1 protein.

GN T06D8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;

RN SEQUENCE FROM N.A.
RA Palmer S.;
RL Submitted (Apr-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RT none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RT Science 282:2012-2018 (1998).
RL EMBL: 249130; CAAB8964.1; -
DR InterPro; IPR003014; PAN.
DR Pfam; PF00024; PAN; 3.
SQ SEQUENCE 1829 AA; 187218 MW; 18A4564A71B3C3E CRC64;

Query Match 7.3%; Score 585.5; DB 5; Length 1829;
Best Local Similarity 23.2%; Pred. No. 1.5e-15;
Matches 375; Conservative 174; Mismatches 662; Indels 405; Gaps 65;

QY 2 GAGSSTE-----QRSPQAGSDTPSELVLSGHGPAEASG-----AAGDPADADPAT-- 49
Db 106 GGEETTVAAVAVSSGEEPPASSSTSVPTLSKDDQVTEASGEETTTAAATEASSETTTS 165
QY 50 KLPQKNGQISVNGVAE-----QGDVHVOENOGGEE-----VYDDVQ 91
Db 166 AYTEGSEETTTSAVTEASSEATTPAGTEASGEETTTSAVTEGSEETTVAAVAVSSGE 225
QY 92 RSE-----DYREKRVEMANSTAVEDIRKQGEETSEIIQIPASENNVEEMV-- 142
Db 226 EPASSSTSIPTLSKDDQVTEASGEETTTAAATEASEETTTSAVTEGSGEDTTVAVAVEL 285
QY 143 ---QPAESQANDVGFKKVFKFTVKKDKNEKSDTVQLTVKKDEGGAASVAG- 198
Db 286 SGEPASS-----TSIPTLSKDDQV-----TEASGEETTTAAAT 321
QY 199 DHOEPEVETAVGESASKESEIKOSTEKQGTLSKQESSTEIPIQASDQAEAEADDEB 258
Db 322 EASEETTTAVTEGSEETTVAAVAVSSGE--EPASSSTSIPTLSKDDQVTEAS--GE 376
QY 259 EKOGEKPTK-SPESSPVP---NSFTTSFKKFTTGAMGRKTKTSFKSKEDDLETAEK 314
Db 377 ETTTAAATEASEETTTSAVTEGSGEDTTVAVAVSSGEPASSSTSIPTLSKDDQVTEA 436
QY 315 RKEQAEKVDEEKEKTEPASEEQEPAEDT-----DQALADYKEVLEPLEDOV 364
Db 437 SGEETTTAAATEASEETTTSAVTEGSGEDTTVAVAVSSGEPASSSTSIPTLSKDDQV 496
QY 365 GDLEASSEKCAPLATEVPEDEKMAHOEVAAVHVSTVEKTEEEQGGGGAEGGVVVEG 424
Db 497 --TEASGEETTTAAATEASE-----TTTSAVTE--GSGEETTVAAVAVSS 538
QY 425 GSSLPEKTLAPPEVQOEAPEAEILKMSREMCVSGDHT---QLTLSPEKTLPGHPEG 481
Db 539 GE---EPASSSTSIPT-----TELSKDDKVTEASGEETTTAAATDASSSETTTSAVTEG 588

QY 482 IYSEVEMLSQGERIKVQSGPLKTLFSSSGGLKKGKQKGRGGGDEPGEYO-HITTE 540
Db 589 SGEETTVAAVE-----SSDEPPASSSTSIPT 616
QY 541 SPESADEQKGESS-----ASSPEEETTCLEKPLEAPODGEAEGETT-----SDGEK 589
Db 617 L--SKDDQVTEASGEETTTAAATEASEETT-----TSVTEGSEETTVAAVAVSSGEE 668
QY 590 KREGITTPASFQKAVTPKRRVRRPESDKEELEKVKATLSIDSTVSEMDEKVTGVE 649
Db 669 -----PASSSTSIPT-----ELSKDDKVTEASGEETTTAAATDASSSETTTSAVTEGS 716
QY 650 -----EQKEPEPKRVDTSVSWALLCVGSSKKRAKASSDDEGPRTLGSGSH 699
Db 717 GEETTVAAVAVSSSDEPPASS-STSIPTEL-----SKDDQVTEASGEETT 759
QY 700 RAEEASDKKEAGTDVAVPASTOEDQAQ-----SSPEPAGSPSEGGVSTWESFKRLVTPR 756
Db 760 TAAATEASEETTTSAVTEGSEETTVAAVAVSSGEEPPASSSTTS-----IPT 805
QY 757 KSKSKLEEKAEKEDSSVQLSTEIEPSREESVWSIKKIPGRKKRADGKQOATVEDSGP 816
Db 806 ELSKD--DQVTEASGEETTTAAATEASEETTS-----AVTEGSG- 843
QY 817 VEINEDDPNVAVPVLSEYNAVEREKMAQGNTELPOLGAVYSEELSKTLVHTVSAV 876
Db 844 -----EDTTVAVAVSS-----GEOPASSSTSIPT-----TELSK----- 872
QY 877 IDGTRAVTSVEERSPSWISASVTEPLEHTAGAMPVEEYTKDIIAEETPVLTOGLEG 936
Db 873 -DDQVTEASGEETT-----TAAATEASEETTTSA-----VTEGS--GETTVAAVAVSSG 919
QY 937 KD-----AHDVMTSEVDFTSEAVTATE-TSEALRTEEVTEASGAETTTDWS 983
Db 920 EPASSSTSIPTLSKDDQVTEASGEETTTAAATEASEETTTSAVTEGSGEETTTSAVT 979
QY 984 AVSQTLDSPDTTEKTPQVQVSVGLDTEEBERQALQAVADKVK---ESQVPATQ 1039
Db 980 EGS-----GEEITTSVAP-----EGENSTTEAPVTVGSIEIIPSSSESSSTTTH 1024
QY 1040 -----TVORTGKALEKVEEVEDSEVLASEKEDVMPKCPVOEAGAEH-----LAQ 1086
Db 1025 DPSIVITPKREVSVSTIERVWS-KTSSEBAEKKIIGHQTKDDADKEDBNMPAVT 1083
QY 1087 GSETQATPESLEVEPVTAADVHVATCOVYIKLQOLMEQAVAPESSEVLTDSFTNGSTPLA 1146
Db 1084 ANPAGSTTESAENVTSIGEED-----ENIKWAKELGKOPAD-----LA 1123
QY 1147 DSDTADGTCQOQDITISQPSKATPAARQSYTBEAATAQKEPSTLPNNVPAQOEHNGEER 1206
Db 1124 KLAARDGVNLTETPAKDSGETAHAHVEDQVSTSESSISGEETTTTV--NKETTEHHENS 1181
QY 1207 GRDVLPTQOELTAAAVPLAKTEVGOGEVMDLGEKVEKOEQVFNVSQPSQKADVT 1266
Db 1182 G-----EEBDAPAFYTGAFPTDSTTEASVGTSA--ITDPTTSVAADSESTISAGVQ 1231
QY 1267 YDSEWVGA--GCOEKESTEVQSLLEGEMETDVEKEKRETKRPQVSEEGQETPAPEH 1324
Db 1232 SSSAIIISATVASEBOTSEATSVIESGE--EVTTDENLVSTVAOLEGSGGATIAESK 1290
QY 1325 EGTGKPVLTIDMPSESKKALGSLGSPS--LPQODRAGCTIEVQVQSLD-----TTVQY 1378
Db 1291 D-----EDSVTTEATSQSTTVSSSDSGSGEYVAPADSESTTSSQSTTDESGVTAES 1346
QY 1379 AAAYEKVLE--TVVISGEGSPGCGAHLPAEKSSATGCHTLOHAEDTV---PLGPES 1433
Db 1347 KDEBSESTTAPAFVTSKTSGSE-----DEEDSPDTHFLTGIDETMKNKSLVPT 1397
QY 1434 QAESIPITVTPAESTLHPDLQGEISASQRESEBEDKPADPADGKESTAIERY 1489
Db 1398 HRDLPNNVGFPVPSSEPKRKNPDE-----EEEBEEDDQTKSDVDVENVSKKI 1444

RESULT 10

```

076891
ID 076891 PRELIMINARY; PRT; 5327 AA.
AC 076891;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE EG:49E4.1 protein.
GN FUTSCH OR EG:49E4.1 OR CG3064.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Papagiannakis G., Spanos L., Siden-Kiamos I., Louis C.;
RT "Sequencing the distal x chromosome of Drosophila melanogaster.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Benos P.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031128; CA20006.1; -.
DR FlyBase; FBgn0015390; futsch.
SQ SEQUENCE 5327 AA; 575942 MW; FEFER23A118FF38A CRC64;

Query Match 7.1%; Score 574; DB 5; Length 5327;
Best Local Similarity 21.6%; Pred. No. 1.5e-14;
Matches 415; Conservative 285; Mismatches 734; Indels 484; Gaps 82;

QY 6 STQRSPQAGSDTSELVLSHGPAEASGAAGPAD-----ADPATKLPQKNGQL 58
DB 2744 SVVESKDDAEKESRPESVIAEGEPVPRESKGPLSDKTSRPGSVVESVTABDEKSEQ 2803
QY 59 SSVNGVAE--QGVHVQENQEQEEVVDVQGESEDVREKRVENA-----ANSTA 112
DB 2804 SRRESVAESVKADTKDGKQGEASRPSSVDLL---KDDKEQESRRQSTGSHKAWTM 2860
QY 113 VEDITKGQSETSEIIEQIPASNNVEMVQ-----PAESQANDV--- 152
DB 2861 GDESPMDKADKSE----PSRPESVAESIKHENTKDEESPLGSRDRSVAESIKSDITKG 2915
QY 153 -----GFKVFKFVGKFTVKDKNEKSDTVQLLTVKKDEGEAEASVGADHOEP---- 203
DB 2916 EKSPLPSEKVSREPVSVVGSIKDKAESRRSVAESVKPESKDATSAPPSKHSRPESVL 2975
QY 204 -----SVETAVGESAKSESELKOSTE-----KQEGTLKQEOSSTEIPL 241
DB 2976 GSKLDGDKTTTSRRVSADSIKDEKLLVSGEASRPSEAEBSLKDAAPSOETSRPESVT 3035
QY 242 QABSD-----QAAEEAEKDEGEKEPTKSPESPSPVNSSETTSSPKKFF 288
DB 3036 ESVDKDGKSPVASKEASRPASVAENAKDSADEKQRPESLPQSKAGSIKDEKSP----- 3089
QY 289 THGAWGRKXTSPFKSKEDD-----LETAERKEQ----- 318
DB 3090 -----LASKDEAEKSEKESRRSVAEQPLVPSKEVSRPASVAESVKDEAEKESPLM 3143
QY 319 -----FAEKVDEEK-----EKTE-PASEEQPAEDTDOARLSADYEVK 356
DB 3144 SKEASRPASVAGSVKDEAEKSEKESRRSVAESKPLPSKEASRPASVAESVKDEADKSK 3203
QY 357 ELPLEPQVGDLEASBEKCAP--LATEVFDE---LWEAHQEVVAHV----- 398
DB 3204 ESRRESGAESKPLASKEASRPASVAESIKDEAEKSEKESRRSVAESKPLPSKEASRPRTS 3263
QY 399 -----VSTVEKTEEGGGGEAGGVVVEGTGSLPPEKLABEQVPOEAPAEELMKSR 453
DB 3264 VAKSVKDEAEKSEKESRRSDSAESKPL--ASKASRPASVAE--SVQDEAEKSK--ESR 3317
QY 454 EMCVSGGDHTQLTDLSPPEKTLPKH-PEGIVSEVEMLSQERIK--VQGSPLKPLFSS-- 508

```

```

QY 1406 LPAKSSANGHWLQHADTVPLPESQASIPITVTPAPESTLHPDLQSHISASORE- 1464
DB 4372 FESVSTSVTKS--TVLSSSTVQLREBSESTESL-----SSSLKVE-----DSRRRS 4416
QY 1465 -RSEEDKPDADGKSTAIKELKAPPELLLESKSNKIVLNVOTAVDQAFATE- 1522
DB 4417 LSSLAEKOGITNTSLKEDTS-ASASQJELLVQSESSSSISVSEIOTISIAQSNKEI 4475
QY 1523 -TAPETHAVDSQTVQVACPLDSREPNRCWTMMKADAMKHPVQOPREDLOVTLVLEAMA 1579
DB 4476 KQARET-----KVTSQFTTTTSATKDSLKETVLEAFLETETKIVSAKEAFS 4521

RESULT 11
Q9FND5 PRELIMINARY; PRT; 2910 AA.
ID Q9FND5 01-MAR-2001 (Tremblrel. 16, Created)
AC Q9FND5 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Similarity to heat shock protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsie.
OC NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98069011; PubMed=9405937;
RA Kocani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned P1 clones.";
RL DNA Res. 4:291-300(1997).
DR EMBL; AB006702; BAB11602.1; -.
DR EMBL; AB017062; BAB11602.1; JOINED.
DR Heat shock.
SQ SEQUENCE 2910 AA; 325351 MW; A847EC3FE1427DF7 CRC64;

Query Match 7.0%; Score 564.5; DB 10; Length 2910;
Best Local Similarity 21.6%; Pred. No. 1.8e-14;
Matches 404; Conservative 275; Mismatches 645; Indels 549; Gaps 84;
QY 65 AEOG--DVHVOENOG--QEEVVDVDEVDGORES-----EDVREK--DRVERMANSTA 112
DB 319 AEOGSEVTVDEKEEDITQNIIEVQSPSVNESPTIOGEDIESKSLDHEEM----- 372
QY 113 VEDITKQGEETSEIIEQIPA--SENNVEE-----MYQPAESQANDVGFKKVFGVF 163
DB 373 -DKITQDTBOGHVLRVDPVPOSEFLVTEAKTAEFTFSVQEAAILKTININESAHSAG 431
QY 164 K-----FTVKDK-NEKSDTVQLTVKDBEGGAASVGA-GDHOE----- 202
DB 432 EEDGQETKENTEPSPKDLKDKQEDSEITVKTI-ISSDEVRSDDVQAEVGEHTEPCSSSI 490
QY 203 -----PSVETAV-GESASKESLQSTOKSTOKSTOKSTOKSTOKSTOKSTOKSTOK 242
DB 491 KODRHRDRDSIEVKAKETGEIETAVDGSVNH-----IETREVLLEAKREBEKEIKID 546
QY 243 AEDDAAEBADEG-----EKO--EKEPTKSPSPSPSPVNSETTSPFKKFFTHGMAGM 295
DB 547 EEPFLAIEKAETENKVIIEPEIYNNBETSVESESLEKNAEPEAVAN--SDGTEQI 604
QY 256 RKTSTKSKSKEDL-----ETAEKKEQA-----EKVDE----- 325
DB 605 SEEVTVDRKEDIDIAKTEIEIOPSPESKASLEPKKEVHISNTEHEHVLERDVOQE 664
QY 326 -----EKEKTEPA--SEOEPAE-DTDOARLSADYEKVELPLEDOVGDLSEASSEK 374
DB 665 TIESEAVETKEDTQPSLDLKEDEKETEAEETFKTVFSSDEVRSASAVQEDQF-----EH 717

```

```

QY 375 CAPLATEVEFDE---KMA-----HOEVVAEVH-----VSTVEKTEEEQGGGGAEGVAV 421
DB 718 TEPCSSIEKDSHGKSEVEVKSQETVODENTEDKHDLKVPSTSEKYOQNEPE-TLVV 776
QY 422 EGTG----- 425
DB 777 SNTGSYEKSEKSPSLVAVNDKBEINDKINVDVQDTQIMEEPIGLDSNGAQAQIDON 836
QY 426 -----ESLPPEKLABPOEVPOEA-----BPAELMKSR 453
DB 837 ITNETEELVAKPVSLIDPVKSVEQMQKPLSPSEVSESTKTVDEKIEKPEEEVTLVYQ 896
QY 454 EMCVSGGDHTQUTDLSPEEKLPHKPEGIVEVEMLSQERIKVQGSGLK--LFSSGL 511
DB 897 EGQVDS-----YGLETEETV-SVPESI--ELEEQPQESVVIDPTLPLOKPTLESSPEV 948
QY 512 KKLGGKQKRGKGGGDEPEEGYHIHTESPESADEQK-----SSASSPPEEPT 564
DB 949 LBSKTV-----DEKIEKTDSTELGELIAQESVTDLTPLQESSOPNEQEKET 999
QY 565 CLK--GPLEAPQDGAEEGTTSDGKKREGITPWASFKNVTPKGRVRRPSDDKEEL 622
DB 1000 KLEKHEPTNEVKSDEVIEVLASPSKLEGET-----VVEAKENIKENEEQQA 1051
QY 623 EKVKA-----TLSTDTV--SEMQDEVKYGEQKPEEKKRV-----DTSVW 666
DB 1052 EKIQKSLTVQVESPSLSFSSBEOQHV-TVAEETVDEKAKEEVPMLQIKNEDATKIH 1110
QY 667 EALICVSSKKRARKASSDDEGGPRTLGDSHRAEASKKEAGTDAVPASTORODAO 726
DB 1111 ETRV-----EAKRIGSLTE-----ISINQOEBOVKE-----ACKSEQKE 1151
QY 727 GSSSE-----PAGSPBEGSVSTWSEPKLVTPRKKSCLKLEAKDSVEQLST 777
DB 1152 ISTNENINVENTYALHSVEAAEEFATNGESLIDVET---TKSVLIRKKEEBAEMKT 1207
QY 778 EIEPS-----RESWVSIKKFIPGRKKRADGK--QEOATVEDSPVE-INEDDNPVAV 829
DB 1208 DAEPLDLAIEKEELETVTV-----QDAIVNNEETTAHESLSLGDNHQENAEVP 1260
QY 830 VPLSEYNAVEREKMEAOQNTLPLQLAGAVVSELSK-----TLVHTSVAVIDGTRAVT 884
DB 1261 EKTQULDDAEOISREVVYDTE-----READITKEIKEQEGTVLETETIQED-IESET 1314
QY 885 SYEERSPSNISASVTEPLEHTAGAMPVE-----EVTKEOI 921
DB 1315 SLEKKEVDQSSKQTEHEHVLERDIPQCEITKABAVDTSTVEBAAILKTLETNISPEEA 1374
QY 922 IAEETPVLQTLPEKDAHDDMTSEVDFTSAVAT-----ETSEALRTEEVTEASG 974
DB 1375 MHESTSLDKVDKEQKEA--TVKTVIFSNVGTSDAQAEFGHTPECSSEIIDEQSG 1431
QY 975 AETTDM-----VSASVOLTSPPTTEAT 999
DB 1433 SEESVYVSKETVQGESSEKQVNMULDVQSGSEKYOENEPDLSLVK--TEKGDVFEELP 1490
QY 1000 PVQE-----VESGV--LDTEEBERQTAIIQAVADVKAYESQVATQTV 1041
DB 1491 SVVEGAGLDETHNQTLILDVSEVVKQSLDTPBEETSTKIDKIDKREB--VTLHQG 1548
QY 1042 QRTGSKALEKVEV-----BEDSEVLASEKQVMPGAPVQOAGAEHLAQOS 1088
DB 1549 REEGSYGLDTQDEAVSVLESRELQPOOEELCLANBENETKQO---EEQVDKHEPTKE 1605
QY 1089 EFG--QATPESLEVEPVADVDHVAICVILKQQLQMEQAVAPESSETLIDSTNGSTPLA 1146
DB 1606 EVSNQOSP-----VEISNEVLQVSSASI-----SEGE-YETVYEAKEIGBOYA 1651
QY 1147 DSDTADGTQODETIDSDSKATPAAVRQSQVTEEAATQKEEPTLPNNVPAQEEHGE-- 1204
DB 1651 DK-IQSPFETGIVAHASHLSPSSSEKHEHETVSEKTDDEK-----VKQAEPIGDWR 1701

```

QY 1205 EPCRDVLEPT-----QOELTAAAVPVL-----KTEVGOGEVDWLDGEKVEE 1248
 Db 1702 ERGLDIAETTHLSLPVDQEDVDEIHPVALPLDQEKVTSTKGETKSSEAEDDKPD 1761
 QY 1249 QEVFVHSGPN-SQKAAADVTDSEVMGVAGCQEKESTEVQSLSLEEGSMETDVEKEKRETK 1307
 Db 1762 EHVDSSTSPMLSEKNDNETQS-----KTSADVCMQOESGTLVPKPEESKEDK 1811
 QY 1308 PEQVSEGEGETAAPHHEGYTKPVLTLDPSSRGKALGSL-----GGSP-----SLPD 1357
 Db 1812 SOBISETIEBIATSDQ-----TLPIETSHDNTLSSELVSEQDDQSPKKVEBIHE 1862
 QY 1358 ODRAGCIEVQVSLDITVTOTAEAVEKVIETVSETGE-----SPECVGAHLLPAEKSS 1412
 Db 1863 EEPKEAHVDATSERNLPLVTSDA-DNTLSSQLVSETKEBHKQAGSILPTEIIPRESSD 1921
 QY 1413 ATGCHWTLQHAEDTVPLGPSPQAESIPILVTPAPESTLHPDLOCEISAS-QRERSEBEDK 1471
 Db 1922 EALVMSLASREDDKVALQEDCADDV-----RETN---DIQERSISVETESVGETK 1971
 QY 1472 P-----DAGPDADQKESATAIEKVLKAEPEILELSKSNKIVNLVIQTAV-----DQFARTE 1522
 Db 1972 PKHEDEIRDAHVETPTAPIILEENDSETLIAEAKGNBEEINERTVALDHBEFVNH 2031
 QY 1523 TAPETHAYDSQTO 1535
 Db 2032 APKLEETKDEKSO 2044

RESULT 12

Q9NFS3 PRELIMINARY; PRT; 16215 AA.
 AC Q9NFS3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE D-titin.
 GN SLS OR D-TITIN OR CG1915.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]

SEQUENCE FROM N.A.

RA Zhang Y.Q., Broadie K.S.;
 RT "Characterization of Drosophila D-titin gene";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL; AJ271740; CAB93524.1; -.
 DR HSSP; P56276; 1TLK.
 DR FlyBase; FBgn0003432; sls.
 DR InterPro; IPR002106; AATRNA_ligaselI.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003600; Ig_Like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00041; fn3; 5.
 DR Pfam; PF00047; ig; 50.
 DR Pfam; PF00018; SH3; 1.
 DR SMART; SM00408; Igc2; 15.
 DR SMART; SM00410; IG_Like; 34.
 DR SMART; SM00326; SH3; 1.
 DR PROSITE; PS00179; AA_TRNA_LIGASE_II_1; UNKNOWN_1.
 DR PROSITE; PS00002; SH3; 1.
 DR KW Immunoglobulin domain; SH3 domain.
 SQ SEQUENCE 16215 AA; 1841509 MW; 242C8765E00F7603 CRC64;

Query Match

Best Local Similarity 6.9%; Score 558; DB 5; Length 16215;

Matches 390; Conservative 303; Mismatches 720; Indels 490; Gaps 79;

QY 7 TEORSPEQAG-----SDTP-SLVLSCGHPAAEASGAAGDPADADPATKLPOK 54
 Db 12422 TTKRMLRFPAGEGEIIEIIVRDDQPEAEIITIVVEYEP-----EPVNDQEKPEPK 12472
 QY 55 NGQLSVNGVAEGGDVH--VQ-----EENQOGEQEBVDE-----DVGORE 93
 Db 12473 KTR-----KVKDDIHDYIOKLIETPKTELEKYEKIEFPIVVKDKPLDSIDVLDES 12526
 QY 94 SEDVREKDRVEEMAANSTAVDITKDGQEBTSEIIEQIPASENNVEEMVQPAESQANDVG 153
 Db 12527 PKEVQKDKKSRITICKQCQVEEAPQEPPIPVQILEVKPV-EVDVKEVITDCQPKVQEK 12585
 QY 154 FKVPFVFG-----PKFTVKDKNKSQVTLTKKDEGECAEASVAGDHOEVSFVET 207
 Db 12586 TRKVLKGIPEEGTTFKITM--IESEDNSV--TVIVDEEPIASQSTEEHPEOSKEK 12640
 QY 208 AVGESAS-----KESL-----KOSTEKQEGTLKQEQSSTSEIPLQAESDOAA 249
 Db 12641 LAPKPKTVRKVKKDDLSDYVKLIEEIPKVDLEKVE---KVMPEKPKVCLTV-SDSIP 12696
 QY 250 EBEAKDEGE-----BKQEKPTKSPSPSPVNSSETTSFKKFFTHGWMGRKKTSPFK 303
 Db 12697 EEPKPKSQPISVLPTDTTKPKTKTKPTKPTEDTDOQVPDEPTTTVDITDIPELPTQT 12756
 QY 304 SKEDDLETAERKQEAQVDEEE-----KKTPEASEEQEPAEDTDQARLSADY 353
 Db 12757 AQPEDTATAQITPSAQEEKSTODDTTKDTTQKTVKHKTKP--DTQKSVETSELPEVHKDY 12814
 QY 354 -----EKVELPLEDOVGLEASSEKCAPLATEVEDE----- 385
 Db 12815 QISIHEELVEEQPEKILIEVRVID-EVAEVEESQPIVEVEDEEPEQATEETVEDVTKP 12873
 QY 386 -----KMEAHQEVVAEVHVSTVEKTEBEQGGGGAEGGVVVEGTGESLPPEKLA 435
 Db 12874 KSKKKVKKVKKDDHDLIKKLEQIEKTELEKYEKIEFDPVKPLKPEFAALEPIKIER 12933
 QY 436 PQEVP-----QEAPEABEL-----MKREMCVSGGDHTQLT 466
 Db 12934 KEQKPTKVTILDATDVPKTKLPSKRKPKABELTVQLPKFLKARMLVE----- 12985
 QY 467 DLSPEKTPKPE-CIVSEVEMLS-----SQERIKVQSPLKLFSSSGLKSLGKKQK 520
 Db 12986 --YPPAPLIPKTTDIGAINDGELSNIAGEABELFKPHKTKKI-----KKIK 13032
 QY 521 GKRGGGGBEPGEYQIHHTESPESADEQKSGSSAPPEPETTCLEKGLPEAPOGEAE 580
 Db 13033 -----DDLEKVELEKYEKVISSEEPKPYKK-PEKAPKEEKQ 13072
 QY 581 EGT---TSDGEKK-REGITPWASFKKQVTPKKVRRPSESDKE-----EELEKYSAT 629
 Db 13073 EDVKLKLGGKKKKPEEAP-----ENVTLKNIPQKQPEVEEVELKQKPEVEIEEQT 13127
 QY 630 LSTDSSTV-----SE-----MQDEVKTVGEQKPEEPKRRVDSVNEALICVSSSK 676
 Db 13128 KKPQGEFVVEPEPSEFDRPEYVPDELSQIEHPKPEKPKPSKT-----K 13174
 QY 677 KRARKASSDDGGPRTLGGDSHRAEASKDKAEAGTDAVPASTQEQDQAGSSSPSPAGS 736
 Db 13175 YRPKDKSKSEPETIIVEIIVAGVPKEEALPEQDVKPKERDAPEETDSEIKLRPQAS 13234
 QY 737 PSEGEQVSWESFKRLVTPRKSKSKLEKAEADSSVE-----QLSTIEPSRESWSVS 789
 Db 13235 KDN-----PDEQALVTP--KAEPIPOEIEDKADDEKPKKSKPKKVPQKPEQ-- 13282
 QY 790 IKKFIPIRRKKRADGQEQATVEDSGPVINE-DDPNV-----PAVVPLEYNAVEKEKME 844
 Db 13283 IAKBED--EEFEVSVKEEALVDK--PIEIKPKDVKKVEKPKPEAPSEVVVIEEKP 13338
 QY 845 AQNTLELP--OLLGAVYVSELSKTLVHTSVAVIDGTRAVTS--VEER----- 889
 Db 13339 EEPSEIPVEYKLTITVLEPEADPKE--HOVKVIDFERQETTEEVEIEEKVTRKKKPKP 13396
 QY 890 -SPSWISASVTEPLEH-----TAGEMPVPEVTEKDIIAETPVLTTQTLPEGKDAHDD 942

```

DB 13397 QCPPEFEVTLKPKPEQIOIPDVSAISLPIEPPEQKPPQYEVNLTQTTTBE--EPNUV 13454
QY 943 WVTSEVDFTSNAVATATSETSEALRTEBTEVTSAGAEFTTDMVSAVSQ--LTDSPTTE-EAT 999
DB 13455 QLAIVEKVTKEVKKVKKEDKIVVVEAEKQPEETIVEVEKQEKKEKSEKPSYEFKIS 13514
QY 1000 PVOQEVESGLDTEEF--EKQTAIILQAVADK-----VKE-----ESQVRA 1037
DB 13515 ETQSIIEKKIIEVAEAEAEETPKVKEKVAEKDYSFTLKEDEEKVITVDDQPEEAP- 13573
QY 1038 TOTVORTSGKALEKVE-----VEEDSEVLASEKEDVMPKQVQEAENHQAQS 1088
DB 13574 VEVEFKKKKEPEBAVAEAVMTPEKIVETSVETAKQKTKKPKDEEQAQAIIVVSS 13633
QY 1089 ETQCATPESLEVEPVATDVAVATCOVILQIMEQAVAPESSETLT---DSETNGSTP 1144
DB 13634 EAPVAEEVSEAPE-----SKIVE-----EVIAEKKPEFTIVSESEPKPEBP 13678
QY 1145 LADSDTADCTQODET-----ISQDSKATRAVQ-----SOYTEEATATQK 1186
DB 13679 SVEQFTVKKKRKSVPFADPATEIVKESKPAEVNTEDAHIKTKPKKVTQVEAEELKI 13738
QY 1187 BEPSTLPNNVPAQOEHEGPEGRDVLPTQOELTAAVPIAAKT-EVQGEVDMLDGKKY 1245
DB 13739 KITTEVPQEIPLIEEVSSE--EVITETK---TAPVBEKTYKIGIKETPEKPAEAI 13791
QY 1246 KEQOEVFVHSGNSQKADVTYDSEVGVAGCOEKESTEVOSLSLEGEMETDVEKEKE 1305
DB 13792 VEEBEVVTVEPIEAPKPEVPEEHKRVL-----EETPRELVEEVLIEEVKVIIRK 13842
QY 1306 TYPEOVSEGEOTAPAEHEGTGKAVLTLDMPSSRGKALGSGSPSLPDDOKAGCIE 1365
DB 13843 KKPKEIKEPKPEABVTV---STPKPV-----BEVATSSIAVTPQPTPEEPAADK 13889
QY 1366 VOVSJDTVTQTAABAVKVIETVNSE-----TGSEPECCGAHLPAKS 1411
DB 13890 ITTIEEBE---TPPOELVQEEIEEIEBPKAPEQPTDTPTATKSEKPTVEELPEBOV 13946
QY 1412 SATGCHMTLOHAEDTVPLG---PESQAESIPITVTPAPESTLHPDQGEISASQERSEB 1468
DB 13947 -----TIQKKKKKAPVEVEVEPEAPFLVKRKTVOEVT---BEAKITSKKPVKEE 13995
QY 1469 EDKPDAGPADGKESTAIEKVLKAPETIEL-----ESKSNKIVLVNITQVAVQPAR 1520
DB 13996 E-----AAELKVTITEEIPTEPEVOEIIIEIEIEEKEPAEYVIEKESQPAVED 14047
QY 1521 TET-----APETHAVDSQ--TQVPACRLDSREPNRCWTQKMDAKM----- 1558
DB 14048 KEVSLPKKKPKAPVIEEPEEITLKKPKVSEEVQBEAKIVKPKKIDEVAVADLTQKV 14107
QY 1559 -KHPVQPREDLQVTLVLEAMQPRKCLPR-----LQDKAP 1593
DB 14108 EEEVVEPEPIVEEVEIEFEFKKKPKPEPEDIVAIIVLKKP 14150

RESULT 13
Q917U4 PRELIMINARY: PRT; 6815 AA.
AC Q917U4:
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DE 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE CG18242 protein.
CN CG18242.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OK NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;

```

```

RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George K.L., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abtill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Butris J.M., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekhailov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiter K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC - SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AE003473; AAC22226.1; -.
DR HSSP; P56276; ITLK.
DR FlyBase; FBgn0035301; CG18242.
DR InterPro; IPR003862; FNIII_repeat.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003600; IG_1like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00041; fn3; 5.
DR Pfam; PF00047; ig; 11.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SMO0060; FN3; 2.
DR SMART; SMO0408; IG_C2; 5.
DR SMART; SMO0410; IG_1like; 6.
DR SMART; SMO0326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW Immunoglobulin domain; Repeat; SH3 domain.
SQ SEQUENCE 6815 AA; 779559 MW; A4E244001A4EBA01 CRC64;

```

```

Query Match 6.9%; Score 557.5; DB 5; Length 6815;
Best Local Similarity 20.5%; Pred. No. 8.9e-14;
Matches 393; Conservative 304; Mismatches 716; Indels 503; Gaps 80;

```

```

QY 7 TEQSPPEQAG-----SDTP-SELVSGSGPAEASGAAGDADNDPARKLPK 54
DB 2989 TTKRMKLRPPAGGEIIEIIVARDPQPEAEITVTEP-----EPVNDKPEPK 3039
QY 55 NQGLSSVGVAEQGVH--VQ-----EENQGEVEVDE-----DVQGRE 93
DB 3040 KTR-----KVKKDDIHVYQCLIELTPKTELEKYEKTEFPPIVKDKRLDSPIDVLES 3093

```

Qy	94	SEDVREKDR	-----VEEMAANSTAVEDITKDGQEBETSEIIIEQIPASENNVEE	140
Db	3094	PKYQKDKKRSRTKVPNBETPVQEQYAKVNVVBEAEPQEIPIVQLIEVKPV	-----EVDVPE	3152
Qy	141	MVQPAESQANDVGFKKVFFVG	-----PKFTVKDKKNEKSDTVQLLTVKKDCEGEGAEAS	194
Db	3153	VITEDGKPVQEKTKTKVLLKIGIPEQTTFKITM	-----IESEDNDSV---TVIVDEEPEIASP	3207
Qy	195	VGAGDHOEPPSVETAVGESAS	-----KESEL-----KOSTEKQBGTLTKQBSQS	236
Db	3208	QSIEEHPQESKEKLA PKPKTKVTKVKODLSYVKKLIEBIPKVDLEKYE	-----KVEMPE	3264
Qy	237	TEIPLQAESDQAABEBAKDEGE	-----EKQKEBTKPSPESSPVNSSETTSSPKKFFTH	290
Db	3265	KPVKLTV-SDSIPEEPKPKSOPISVLPTTKTKTKTKPTKPTEDTDOQVPDPEPTTTV	-----	3323
Qy	291	GWAGWRKTSFKKSDDLLEAKRKEQEAERKVDSEB	-----KEKTEPASEBOEP	340
Db	3324	DTTDDIPELTTQTAQPEDTATAIQTPSAQEKSTQDDTKOTIQTKVHKHKT KP	-----DTQKS	338
Qy	341	AEDTDQARLSADY	-----EKVBLEPQDVGLEASSBEKCAPLATEVDFD	---385
Db	3382	VETSELPEVHKDQI0SIIEHLEVEEOPKILLEVVID	-----EVAEVEESQPIVEVEDEPEQ	3440
Qy	386	-----KMEAHQEVVAHVHSTVEKTEBEEOGGGEABGGVVVE	422	
Db	3441	PATEETVEDVTPKSKKKKVVKKTKTDHDELIKMLEQIEKTELEKIEKIEFDFVPKKLK	3500	
Qy	423	GTGESLPPEKLAEPQVVP	-----QEAAPAEEL-----MKSR	453
Db	3501	PEPAALEPIKIEKEOKTKVTILDATDVPTVKLPSKRKEKFAEELTVOLPKFRLKAR	3560	
Qy	454	EMCVSGGDHQTQLDLSPEKTKLPKHPE	-----GIVSEVEMLS-----SOERIKVQSGPLKCLFS	507
Db	3561	MVLVE	-----YPPAPLIPKTTDGAIKDNGELSRNIEEAELIKFPHKTKKI	---3608
Qy	508	SSGLKKLGGKKKGKGGDEBEPGEYQHIHTESPESADEQKGSASSPEEPETTCLE	567	
Db	3609	-----KKIK-----DDLEKVELEKYEKYSISEBEPEKTPYK	3640	
Qy	568	KGPLEAPDGEAEECT	-----TSQEKK-REGITPWASFKKWTPPKVRVRPSESDEK	---619
Db	3641	K-PEKAPKPEKQEDVKLGLGKKKPKPEEAP	-----ENVTLKNIPOKPOVEEVEELK	3694
Qy	620	-----EELEKVKASATLSSTDSTV	-----SE-----MODEVKTVGEQKPBKPKRRVDTS	663
Db	3695	QKPEVEIIEEQTKPKPDGEFVVEFPESFDFRPYVDELEQIEHPIPEKKVKPSKT	---3753	
Qy	664	VSWEALICVGSKKRKARAKSSDDGGPRTLGGDSHRAEASDKOEAAGTDAVPASTOEQD	723	
Db	3754	-----KYKPKDKSKSEPIVSEIIVAGVPKEEAI	PEQDVKFRKPERDAPET	3801
Qy	724	QAQSSSPGAPSGEGSVTWSFKRLVTPRKSKSLEKEAEADSVE	-----QLS	776
Db	3802	DSEIKLRVPQASKDBN	-----PDEQALVTP--KAEBEPIQEIETDKAIDDEKKPKKASKP	3853
Qy	777	TEIETPREESSWVSIKFIPIRRKRGADKQEOATVEDSGPVEINE	---DDPNV---PAVVP	831
Db	3854	KKVQPKQEQ	-----IAKEEP--FEFESVVKBEALVDK--PIEIEKPDVKVKEKKPEAP	3905
Qy	832	LSEYNAVERKMEBAQNTPEL	---QLLGAVVYVSELSKTLVHTVSVAVIDGTRAVTS--V	886
Db	3906	VSEVWIEEBPKPEEVPBIEPVEYKITTVTLEPDAPKE	---HQVKVDFDERQETTEVI	3963
Qy	887	BER	-----SPSWISASVTEPLEH-----TAGAMPPEVEEVTETKOIIAETTPVL	929
Db	3964	EKKVTRKKKPKPQOPEFVTLKEPFKEQIQPDVVVSAEISLPIEPEEQKPEQVEVELKI	4023	
Qy	930	TQTLPECKOAHDDMWTVSEVDFTS	SEAVTATETSEALRTEBTEVTEASGABETTDMWSAVSQ	---987
Db	4024	TQTTPE	---EPNDVQIIVAEKVTKPKVKVKEDKIVVVEAEEEKQPVETIVEVEKQBEKK	4081
Qy	988	LTDSPDPTTE	EATFPQOEVESGVLDTSEE---ERQTOAILQAVADK-----VKE-----	1031

RT TAC and BAC clones".
 RL DNA Res. 7:217-221 (2000).
 DR EMBL; AP002057; BAB03174.1; -
 SQ SEQUENCE 2081 AA; 232851 MW; D36031E1F85EFFP29 CRC64;
 Query Match 6.4%; Score 517; DB 10; Length 2081;
 Best Local Similarity 19.8%; Pred. No. 9.4e-13;
 Matches 324; Conservative 293; Mismatches 721; Indels 302; Gaps 55;
 Oy 5 SSTEORSPROPAGSDTPSELVLSHGPPAAEASGAADPADADPATKLPOKNQ-----L 58
 Db 191 ASSESSSTHNNVTTGSMNV-ETNGENSETQEKGDVGSNGDVSMENLGNKVDEL 249
 Oy 59 SSYNGVAGQGDVHVQENOGQEEVVDVQGRESEDVREKRVEMANSTAVADITK 118
 Db 250 KEGNNVVENGE--TKENNENVENNEKEVEGQES-----IGSAIE- 290
 Oy 119 DQGEETSEIIEOIPASENNVEEVD-ASQANDVGFKKVFKFVGFYTKDKN----- 172
 Db 291 KULESKEDVKEVEAAKNQSSMTENLGEAQNN-----GVS-TIDNEKEVEGQ 339
 Oy 173 ---EKSDTVQLTVKKEGGAASVAGDHPESVETAVGESASKESELKOSTEKQEGT 229
 Db 340 ESIJEDDIEKULESKEDVSEVEAAKNQSSMTGKLEBAQNNVGTNETMNSENKGSQ- 398
 Oy 230 LKQEGSTETPLQAEESDQAAEEBAKE-----GEEKQEK---EPTKSPESSTPY 276
 Db 399 ---ESTNDKQVNAATTNDEHKKEKEETHENNGESVKGENTLENKAGNESMKENLENKY 455
 Oy 277 NSETTSSFFKFTFHGAGNRKKTSPFKSKEDDELTAKEKEQ-----EAEK 322
 Db 456 GNEELK-----GNAASYAKTNNESSKEEKSORSNEVYMKETTKGENVNIQGS 507
 Oy 323 VDEEEKETEPASEEOBPADTDARLSADYEKVELPLEDOVDLEASSEKCAPLATEV 382
 Db 508 IODSTDNLENKEDVYKPKVDANESDNGSTKERHQ---EAQVNN-GVSTEDKLDNIGAD 563
 Oy 383 PBEKKAHQEVVAEVAHVSTVEKTEEBQGGGGAEGVYVETGESLPPREKLAPQEPVQ 442
 Db 564 EOKKKNKSVAVTTNDGDHTEKEKEETQGNNGS---VKNNLENKEDKELKODESVGAK 620
 Oy 443 ABAEALMSREWCVSGD---HTQLDLSPEBKTLPKHPE---GIVSEVEMLSGERIK 496
 Db 621 TNNETSLSEKREBQTKGHNDJINSKIVDNKGADSNKEKEVHVGSTJDNNESEKEDTK 680
 Oy 497 VOGSPILKTLFSSSGLKKLGSKKQKRGGGGDEPEYQHIHTESPESADEQKGSASAS 556
 Db 681 -----SEVEYKKNQSSSEKGEKENNKDSMEDKTL--ENKESQDSKODKSVDD 728
 Oy 557 PEEPEETCLEKGPLAPDQGEAEBGTSDGE-KKREGITPMASFKKMVTTPKRRV--- 611
 Db 729 KOEBAQI-----YGESKDKDSVEAKGKKES---KENKTKTNNENRHNKEE 773
 Oy 612 ----RSESDKEELEKVSATLSSTDS-----TVSEMODEVKTGEBQKPEPRRYD 661
 Db 774 NVQGNKKESEKVEKKSADKASVETKDNKKLSTENDEAK---EKSGEDNKEDXE 829
 Oy 662 TSVSMALICVSGSKRARKASSDDEGPRTLGDSH----- 699
 Db 830 ESDVDQSV-----EAKKKNMGVDTNVGNKEDSKDLKDRSVEVKANKEESMK 878
 Oy 700 -PAEASKDEACTDAVP--ASTQEBODQAGSSSPREPAPSGEGVSTWESKRLVTPR 756
 Db 879 KRREEVQRDKSSTKEVRPFANMMDIDVQSGESVKKYKDEKKEG--NKEENKQDITNS 916
 Oy 757 KKSXKLEKAEBSVQJSTLEPSEBEBWVSIKKFIPERRKRRADQKQEOQATVEDSG 816
 Db 937 SKQKGDKKKKKK---ESKSNKMKKEEDK---KEYNNLELKQEDNKKETJYSNSKL 989
 Oy 817 VEINERDPNVPAVPLSEVYNAVEREKMEAQNETLPGALGAVVSEELSTLIVHTSVAV 876
 Db 990 KKEKNKKEKKESEDASAKNREKKEVEEKKSKTK-----BEAKKEKKKSGQ---- 1034

Oy 877 IDGTAVTVSVERSPSWISASVTEBLEHTAGRAMPVEEVTEKDIABETPVLTQTLPEG 936
 Db 1035 -DKKKEBKDSEBRK-----SKKEEBSRDLAKKKEETYEK---YESENHKSKKED 1083
 Oy 937 KDAHDMVTSEVDFTSAVTATETSEALRTE-----VTEASGAETTDM 981
 Db 1084 KKEHNDKSMKKEEDKKEKKHBSKRSKKEEDKKDMEKLEBQNSNKKKEDNKEKKSGH 1143
 Oy 982 VSAVSQLDSDPD--TTEATPVQVEVSGVLDTBEERQTOALQAVADKVEESQVPAQ 1039
 Db 1144 VVLVKKESPEKKEKKEKESETEKEIESSSKQKNVEDKKEK-----KSKDQOKKKEK 1195
 Oy 1040 TVQRTGSKALEKVEE-----VEEDSEVLASEKEDVM---PKGPOEAGAEHLAQSE 1089
 Db 1196 EYKSEBEKLLKKNEDRKKQTSVEENKQOKETKEKKNPKRODKKNTTQSGGKKESME 1255
 Oy 1090 TQCATPESLEVEVADVDHATCOVIRLQOLMEQAVAPESSEETLTDEBTNGSTPLASD 1149
 Db 1256 SKEA--ENQKQSQATTQADSDESKNELMQ--ADSQADSHSDQADSDESKNELMQAD 1310
 Oy 1150 TADGTQDETTDSQDSKATPAVROSQVTEBEAATAQKEEPSTLPNNVPAQ---BEHGE 1205
 Db 1311 SQATTQNNNEBRKKQTSVAENKQKETKEB-----KKNPKODKKNTTQSGGKKESME 1365
 Oy 1206 PGADVLEPTQBELTAAPVPLAKTEVQGEV-----DMLDGSKVEEOGVFVHSGPN 1258
 Db 1366 ESKENQKQSQATTQADSDESKNELMQADSQADSHSDQADSDESKNELMQADSQAT 1425
 Oy 1259 SOKAADVTYDSFVMVAGQEQESTEVQSLSEBCEMETVDEKEKRETPQEVSEGE-- 1316
 Db 1426 TQRNNEBRKKQTSVAENKQKETKEKKNPKRODKKNTTQSGGKKESM-BESKAEAEQ 1484
 Oy 1317 ---QETAPHEGTGKPYLTLDMPSSERKALGSLGSPSLPDOKA-GCIEVQVSID 1372
 Db 1485 QKSQATTQGESDESKNELMQADSQADTHANSQ-----DDESKNELMQADSQA 1535
 Oy 1373 TTVQTAEAVERKI---ETVVISFTGESPEVCAGALLPAKSSAAGCHTLOHADIVPL 1429
 Db 1536 DSQTSDESKNELMQADSQADSQT--DSDESKNELMQADSQAKG---ESLBNKYK 1589
 Oy 1430 GPESQAESIPILVTPAPSTLHPDQGEISASQRESEEDPPADGPDAGKESTAIERY 1489
 Db 1590 KGEDNGEV-----GKNSKTIYKVGHBSKDKGTNENGKVEYTERGSDSNIVERN 1643
 Oy 1490 LKAPELLELESKNKIVLN 1509
 Db 1644 GKEKDSIKESDQKTEIN 1663
 RESULT 15
 P91257 PRELIMINARY; PRT; 3484 AA.
 AC P91257;
 DT 01-MAY-1997 (Tremblrel. 03. Created)
 DT 01-OCT-2001 (Tremblrel. 18. Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19. Last annotation update)
 DE Hypothetical 385.7 kDa protein.
 GN F12F3.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID:6239;
 RN (1)
 RP SEQUENCE FROM N. A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N. A.
 RC STRAIN=BRISTOL N2;

RA Fulton B., Wohldmann P.;
RT "The sequence of C. elegans cosmid F12F3.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";

DR Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; U80022; AAC25885.2; -
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR003962; FnlIII repeat.
DR InterPro; IPR003961; FnlIII.
DR InterPro; IPR003599; Ig_c2.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig like.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF00041; fn3; 7.
DR Pfam; PF00047; ig; 10.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SM00060; FN3; 8.
DR SMART; SM00409; IG; 12.
DR SMART; SM00408; IGC2; 8.
DR SMART; SM00410; IG like; 1.

DR PROSITE; PS00225; CRYSTALLIN BETA GAMMA; UNKNOWN_1.
DR PROSITE; PS00142; ZINC PROTEASE; UNKNOWN_1.
DR Hypothetical protein; Immunoglobulin domain; Repeat.
KW SEQUENCE 3484 AA; 385746 MW; 21DD6F6893020F4F CRC64;

Query Match 6.2%; Score 501.5; DB 5; Length 3484;

Best Local Similarity 20.3%; Pred. No. 7e-12;

Matches 333; Conservative 285; Mismatches 560; Indels 461; Gaps 69;

QY 80 QEEVEVDVGQRESDVREKDRVEEANAANSTAVEDITDKGQETSIIIEQIPASENNV- 138

DB 269 RDGLIDGGSDR--MEVRHED---EMRKWLILKQICKD--EEAYACQAI-----NVA 316

QY 139 -----EEMVQPAESQANDVGFKKVFVKFVKTKKDKN-EKSDTVQLLTVKKDREG 190

DB 317 GEAWCFSDVVVHMSERDD-----DKSVDVDDSTVLEEKDDGD-- 357

QY 191 AEASVAGDHOEFSV-ETAVGESAKSESELKQSTEKQEGTLKQ--EQSSTEIPLQAESDQ 247

DB 358 -----DKSKPKTKKIIKKIKETPESEQVTAAPPEQQKISEVDVQSVAEVGAKKKP 409

QY 248 AARE-----BAKDEGEKQEKPTKSPESPSS---PVNSETTSFKFFTHGAWGRKKT 299

DB 410 DAEPDLSAKKDKSKSKSDPEAPESTEKSTTEKPTNDKTS-----KKK 454

QY 300 SFKSKEDDLLETAKRKEQAEKVDEBEKTEPASEQEPAPEDTDQARLSADYKVELP 359

DB 455 AEKTVKPKKEVT--GRPLEAKPVEDKDAQSPSS--SKESSPTDGKKKKQIPKALFIP 511

QY 360 --LEDQVGD-----LEASSEK---CAPLATEVFDEKMAHOEWAHVH 398

DB 512 DEISSREFGDPSTWHSETNITTTIRREGSADAKTPLVEPLASV-SMKVESAKE---KAE 567

QY 399 VSTVEKTEEEQGGGEAGGVVGTGESLPPEKLAEPQ-EVQEAEPAPAEELKMSRECV 457

DB 568 FSKRRSETPDKSRKKEG-----LPPAKSKKDEVTAEKQSTALIESKKEV 617

QY 458 SGGDHTQLDLSPEKT-----LPHKEG-----IVSEVEMLSQBERIKVQSGPLKLF 506

DB 618 ---DEKISQQSDKKNSEVGVPEKAAAPETKQVSEIEV-----PKKTI 663

QY 507 SSGGLKLSGKKKGKGGGDEEPGYQHHTESPESADEQKGESSASSPPEETTC 566

DB 664 KKTEKSDSSISQSNVLKPADDDKSDVDTKSKKTTEDQTKVATDSKLEAADTT-- 721

QY 567 EKGPLEAPOGEAEGETTSGEKKREGITPWASFQKMTVKKRVRRPESDKEELEKVK 626

DB 722 KQIETETVDDKSKKVL---KKTEKSDSFIQKSETPPVPEPTPAESQAQIAEVNK 778

QY 627 SATLSSTDSTVS-EMODEVKTVGEEQKPEPPKRRVDTSVSWEALICVGSKKGAARKASS 685

DB 779 AKQKEVDONLKREAEVAAKKIADKLEKLTAEANIKKTAEEA-----AKQKEK 828

QY 686 DDE-----GGPRTLGGDSHRAEEASKKEAGTGDVAPASTQEQDQAQSSSPPEAGSPS 738

DB 829 DEQLKLETEVSVKKSAAEKLLEKQAOIKKAAEADAVKKQKELNEKNKLEAAKKSAAK 888

QY 739 EGGVSTWESFK---RLVTPRKKSLSKLBK-----AEDSSVQLSTEIPSPSEESVSI 790

DB 889 KLEESAASAKSKVSESVKGEKKTKAGEKTQVSESEPTSKTIDTKDVGATEPADETP 948

QY 791 KKFIPGRRKRGADGEOQATVEDSGPV--EINEDDNNVPAVPLSEYNAV-EREKMEAOQ 847

DB 949 KKKIIKKTEKSDSSISQKSATDSEKVSQKEQDEPTKPAV---SETQMTEADKSKKQK 1005

QY 848 NTELPOLLGAVVYSEELSXTLVHTSVAVIDGTRAVTSVEERSPSWISASVTPELHTAG 907

DB 1006 ETD-----EKLK-----LDAAIAAKTQEAD-----EKSKL 1031

QY 908 EAMPPVEEVTETKIIAETPVLTOTLPEGKDADHDMVTSEVDFTSEAVTATETSEALRTE 967

DB 1032 DAQEKIKKVSDEDDAARKE-----KELND-----KLESEIATKASADKLE 1075

QY 968 EVTEASGAETTDMSAVSGLTDSPTTBEATPVQEVESGLPTE-----EER 1016

DB 1076 EQAQAKKAAE---VEAAKKQKEKDQLK-----LQTEAASKKAAAKLELEK 1119

QY 1017 QTOAILQAVAD-----KVEESQVPAOTQVQ-----RT 1044

DB 1120 QAQIKKAAGADAVKKQKELDKNLEANKKSAAGKLKIEESAASKQVTEQAKLDAOT 1179

QY 1045 GSKALEKVEEVEDSEVLASEKEKVMKGPVQ-----EACAHLAQSSEGTQATPES 1097

DB 1180 KATAKQTKLEKDEKSTESSEKTVDEKPKKVVKKTKESDSSISQKSETSKTVVES 1339

QY 1098 LEVPEVTADVHVATQVILQLOLMQAVAPESSET--LTDS-----ETNGSTPLADSOT 1150

DB 1240 -----AGPSESETQKVADAAARKQKTEKQKLEAEIT 1271

QY 1151 ADGTQOQETIDSDSKATAAAROSQVTEBEAATAQKEEFTLNNVPAQEBHEGEEPRDV 1210

DB 1272 AKKSADEK-----SKLEAESKLKAAEVAEAKKQKEDEQLKLDT---EAASKAAAEK 1322

QY 1211 LEPTQOELTAAAVPVLA-----KTEVGQGEVDWLDGKVKBEQEVFVHSGPNSQK 1261

DB 1323 LELEKQSHIKKAAEVDVAVKKQKELEBKQLESEAAATKKADAELKLEEQ-----KKK 1374

QY 1262 AADVTYDSEVMGVAGQKEKESTEVSLSLEEEMETDVEKEKRETKPEQVSESEGEQETA 1321

DB 1375 AAEIA---LIEIQKEQEKLAQE-QSRLEDEAKKSAEKQKLESETKSKQTEE-----A 1422

QY 1322 PEHEGYTKPVLTLDMPPSERGKALGSLGGSPSLPDODKAGCIEVQVQSLDTTWTQ--- 1377

DB 1423 PKE-----SVDEKPKKKVLKKKTEKSDSSISQSKS 1453

QY 1378 ---TAAVE-----KVIEVTVISETGESPECVGAHLLPAEKSSATGGHWTLOHAEDTVP 1428

DB 1454 AKSTVDAAEETLESDFNLVEKTKVQKVEQSPD-----ESTSATIKRDPAPQKTEISK 1504

QY 1429 LGPESQAESIPITVTPAPESTLHPDLQGISASQERSERSEEDKPDAGPDA----- 1478

DB 1505 QDDGDEKTTTDDGPKPKPDS-----EATPKRVVKKTKQSDSVASDASLADSVKLS 1557

QY 1479 DGKESIAIEKVLK-----APPEILELES-KSNKIVLNIQTAVDQFA 1519

DB 1558 DVVEEKPKKKLKKKTEKSDSVISSTSSVDTIKPESEIPEKAEQWILH-----NRFS 1611

QY 1520 RTETAPET---HAYDSQTQ 1535

DB 1612 -TDSAVESEFPKNAHKDDTE 1629

Fri Dec 13 15:34:18 2002

us-09-902-432-4.rpt

Page 18

Search completed: December 13, 2002, 00:16:01
Job time : 166 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 00:17:16 ; Search time 8528 Seconds
(without alignments)
5446.538 Million cell updates/sec

Title: US-09-902-432-4

Perfect score: 8073

Sequence: 1 MGAGSSTEORSPEQAGSDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US0902432/runat_12122002_141451_25800/app_query.fasta_1.1735
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0_1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0902432 @CGN_1_1_6311 @runat_12122002_141451_25800 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rat.*
36: em.htg.ham.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8073	100.0	6160	10 RNU23146	U23146 Rattus norv
2	8030	99.5	5236	10 RNU41453	U41453 Rattus norv
3	7461	92.4	17854	2 AC115647	AC115647 Rattus no
4	6663	82.5	6195	10 AB020886	AB020886 Mus muscu
5	6663	82.5	6197	10 AF326228	AF326228 Mus muscu
6	6255	77.5	5902	10 AF326230	AF326230 Mus muscu
7	6005	74.4	5134	6 AR071282	AR071282 Sequence
8	4340.5	53.8	6614	6 AX405857	AX405857 Sequence
9	4300.5	53.3	6608	6 AX329766	AX329766 Sequence
10	4300.5	53.3	6608	6 AX410793	AX410793 Sequence
11	4300.5	53.3	6608	9 HSU81607	HSU81607 Sequence
12	4281	53.0	6605	6 AR002603	AR002603 Sequence
13	4231.5	52.4	6297	9 AF001504	AF001504 Homo sapi
14	4097	50.7	104939	9 AL590413	AL590413 Human DNA
15	4097	50.7	171158	2 AC023201	AC023201 Homo sapi
16	4090	50.7	6287	9 AB003476	AB003476 Homo sapi
17	1805	22.4	8439	5 AF308810	AF308810 Xenopus l
18	1585.5	19.6	102822	2 AC097731	AC097731 Rattus no
19	1067	13.2	1027	9 BC000188	BC000188 Homo sapi
20	1067	13.2	1027	9 BC022814	BC022814 Homo sapi
21	807	8.0	160827	3 AC099025	AC099025 Drosophil
22	639.5	7.9	162921	3 AC007853	AC007853 Drosophil
23	639.5	7.9	181132	3 AC008206	AC008206 Drosophil
24	639.5	7.9	227398	3 AE003750	AE003750 Drosophil
25	635	7.9	16377	3 AF202180	AF202180 Plasmodiu
26	634	7.9	193262	3 AC007579	AC007579 Drosophil
27	634	7.9	260367	3 AE003808	AE003808 Drosophil
28	629	7.8	113474	2 AC013932	AC013932 Drosophil
29	612.5	7.6	134855	2 AP004077	AP004077 Oryza sat
30	601.5	7.5	32009	3 CET06D8	249130 Caenorhabdi
31	589.5	7.3	53424	3 AB055861	AB055861 Procamb
32	583.5	7.2	75313	2 AC017682	AC017682 Drosophil
33	582	7.2	144056	3 AC091222	AC091222 Drosophil
34	582	7.2	194006	3 AC010063	AC010063 Drosophil
35	577.5	7.2	39749	3 DMC49E4	AL031128 Drosophil
36	576.5	7.1	179145	9 AC009785	AC009785 Homo sapi
37	576.5	7.1	181360	3 AC104141	AC104141 Drosophil
38	565.5	7.0	302855	3 AE003420	AE003420 Drosophil
39	562.5	7.0	68727	2 AC004426	AC004426 Drosophil
40	557	6.9	71023	2 AC004760	AC004760 Caenorhab
41	547.5	6.8	23554	3 AE003473	AE003473 Drosophil
42	537	6.7	304106	3 DME271740	AJ271740 Drosophil
43	536	6.6	70398	3 AP004150	AP004150 Oryza sat
44	533	6.6	125398	8 AP002057	AP002057 Arabidops
45	517	6.4	73391	8	

ALIGNMENTS

RESULT 1

RNU23146 LOCUS 6160 bp DNA linear ROD 19-OCT-1995
DEFINITION Rattus norvegicus mitogenic regulation ssecks (322) gene, complete cds.
ACCESSION U23146
VERSION U23146.1 GI:1022682
KEYWORDS Rattus norvegicus.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 6160)
AUTHORS Lin, X., Nelson, P.J., Frankfort, B., Tombler, E., Johnson, R. and Gelman, I. H.
TITLE Isolation and characterization of a novel mitogenic regulatory gene, 322, which is transcriptionally suppressed in cells transformed by src and ras
JOURNAL Mol. Cell. Biol. 15 (5), 2754-2762 (1995)
MEDLINE 95257957
PUBMED 7739556
REFERENCE 2 (bases 1 to 6160)
AUTHORS Gelman, I. H.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-1995) Irwin H. Gelman, Microbiology, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029, USA
COMMENT On Oct 19, 1995 this sequence version replaced gi:733449.
FEATURES
source location/Qualifiers
1..6160
/organism="Rattus norvegicus"
/strain="Fisher"
/db_xref="taxon:10116"
/clone="13.2.2"
/cell_line="Rat-6"
/cell_type="fibroblast"
/tissue_type="embryo"
/clone_idb="R1-a (Robert Kraus, Mount Sinai School of Medicine)"
1..32
5'UTR
gene
CDS
33..5044
/gene="322"
33..4823
/gene="322"
/function="possible mitogenic regulatory gene"
/note="src Suppressed C Kinase Substrate"
/citation={1}
/codon_start=1
/evidence=experimental
/product="Ssecks"
/protein_id="AA079517.1"
/db_xref="GI:1022683"
translation="MGAGSSTFORSPEOPAGSDTSEHLVSGHGPAEASGADPADADPATKLPKNGCOLSVNGVABQGDVHVOENQOEQOEVEDVDQEGSESDVREKRDVBEAANSTAVEDITKDGQEBTSEIIEIIPASNNVEMVQPAEQANDYGRKYRKEVGFRTVKRDKNEKSDIVQLTIVKQDBEGEASVAGADHQPSEVETLVGSSASESEVKOSTEKQEGTLKQESSTELPLQASDQAAEEAKDKDEGEKPEKPSSESPVNSSETSPKKEFTTHGMAWKRTSPKFKEDDLTAERKQEKQAEKAVDEBEKEEPA SEBOEPADTDOARLSADYKEVLEPLEOVGDLASSEKCAPLATEVDEMEKHA VVAEYAVSTVETBEREGOGGAGGAGVVEGGLPEKLAPOEVQOAPKULFSS KSRKMCVSGGHTQITLDSPEBKTIPIKPIBGLVSVEMLSQERTKYVGSPLKULFSS SGLKULSGKQKGRKGGGDEPEBYOHITHSPBSAUBQKESASPEEBETLCEKGLPEAPDGAEBGTTSDGKKREGITTPNAPSPKQVTPKRVLRPSDSDEBLEK VKSATLSTDSFVSEMODEVKTVEEQKPEEPKRVDTSPVSEALICVSSSKYARKA SSSDGEGPRTIGDSHRAEASDKBAGTDVPASTOEODAQSSPEPAGSPSEG EGVSTWSEFKULVTPRKSKSKYLEKADDSVQESTLEPSEESMWSTIKKTPGR KRAADGQEQATVEDSGPELVIEDDPNPVAVPVSEYNAVEREKKEAAGNTLPULG AVYSEELSKTLVHTVSAVIDGTRAVTSVEERSWSVSAVTEPLEHTAGAMPVE EYTEKDIIAEBETPVLQTLPEKGAHDMVTSEVPTSEAVTEPLEHTALRTEEVTEA SGAETTMVSAVSQLTDSPTTEBATPTQEVESVLDTBEERQTOALIQAVADKVK ESOVPAOTVORGTSKALEKVEVEEBSVLASKEKDMVKMGVQEGAAHLAGS ETGQATPSESLPEYETADVATQCVIKLQOLMQAIVAPESSELTJDSFNGSTPLA DSDTADGTQDDETTIDSDSKATPAVRQSOVTEEBEATATQKEBPSFLPNNVAQEBHG

BPGRDYLPTQGELETAAPVULAKTEVGEQGEVDMLDGEKVEEBOEVFVHSGPNSOKA
ADPTTYSERWGVAGCGEKESTEVQOSLSLEGEMETVYKRETKREQVSEBGEQETA
APHEHETGKRPVJLIDMPSSERKALGSLGSPSLPDQDAGCIEVQVSLDTVTYT
AAEAEKVIETKPVLSITGESPEVCAGHLLPAEKSSATGAWTQLHAEDVPLGPESQAE
SIDPIVTPAPESTLHPDLOGEISASQERSEBEDKPDAGDADGESTALEKVLAP
EILELESKSNKIVLVNIOFVADQFARTETAPETHAIVDSOQVAPACRDSNEPNRCWTK
MKDAKMHVPVQPREDDQLVTLVEAMNQPRKCLPRQLKAPVSK"
1581..1590
/gene="322"
/note="encodes nuclear localization signal; NLS"
/citation={1}
misc_signal
1797..1811
/gene="322"
/note="encodes nuclear localization signal; NLS"
/citation={1}
misc_signal
2061..2073
/gene="322"
/note="encodes nuclear localization signal; NLS"
/citation={1}
misc_signal
2301..2313
/gene="322"
/note="encodes nuclear localization signal; NLS"
/citation={1}
misc_binding
4230..4358
/gene="322"
/note="encodes zinc-finger domain"
/citation={1}
3'UTR
4816..6160
/bound_molety="DNA, RNA or protein binding domain"
misc_signal
5032..5044
/gene="322"
/note="encodes nuclear localization signal; NLS"
/citation={1}
polyA_signal
5760..5765
polyA_signal
6136..6141
polyA_site
6160
BASE COUNT 1853 a 1409 c 1753 g 1145 t
ORIGIN
Alignment Scores:
Pred. No.: 3,18e-280 Length: 6160
Score: 8073.00 Matches: 1596
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 10
US-09-902-432-4 (1-1596) x RNU23146 (1-6160)
QY 1 MetGlyAlaGlySerSerThrGluInaYrSerProGluInProAlaGlySerAspThr 20
DB 33 ATGGGCGAGGAGGAGTTCCACGACGAGGAGGCCCGGAGCGCGGAGCGACACG 92
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluValaSerGlyValaAlaGly 40
DB 93 CCGACGAGCTGCTGCTCATGAGGCGCATGGCCCGACGCTGAAGCTTCGGAGACACTGGA 152
QY 41 AspProAlaAspAlaAspProAlaThrIyLeuProGluInySasnGlyGluLeuSerSer 60
DB 153 GACCCCGCGAGCGGAGCCCGCCACCAAGCTCCACAGAAAGATGGCAGCTGCTTCT 212
QY 61 ValaAnglyValaAlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGluGln 80
DB 213 GTCAACGGCGGTGCTGGAACAAGAGATGTCATGTCCAAGAGGAAACACAGAGGGGCGAG 272
QY 81 GtugluguValaValaAspGluAspValaGlyGlnaYrSerGluAspValaArgGluLys 100
DB 273 GAGGAGGAAGTGTGTAAGAGATGTGACACGACGATCGAAGAGATGAGAGAAAAA 332
QY 101 AspAryValaGluGluMeAlaAlaAsnSerThrAlaValaGluAspIleThrIySasnGly 120
DB 333 GACCGAGTTGAAGAAATGGCGGCCCAACTCCACAGCTGTGGAAGATATATCAAAAGATGGG 392

QY 121 GlnGluThrSerGluIleIleGluGlnIleProAlaSerGluAenValGluGlu 140
Db 393 CAGAGAGACATCAGAAATAATTGAACAGATCCCTGCTTCAGAAACAAATGTGAAGAA 452
QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
Db 453 ATGGTACAGCCTCTGAGTCCAGGCTAATGATGTTGGCTTCAAGAAAGTATTAAATTT 512
QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
Db 513 GTTGGTTTAAATTCACGGTGAAGAGGATAAAAAATGAAAGTCAAGTACTGTCCAATA 572
QY 181 LeuThrValLysLysAspGluGlyGluGlyValAlaGluAlaSerValGlyAlaGlyAspHis 200
Db 573 CTCACTGTCAAGAAAGATGAAGCGGAGGCGAGAGCCCTCTGTGGAGCTGGAGACCAC 632
QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
Db 633 CAGAGCCCACTGTGGAGACTGCCGTGGAGAGTCAAGATCCAAAGAAAGTGAAGTGAAG 692
QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerThrGluIlePro 240
Db 693 CAATCCACAGAGAAGCAAGAACCCCTGAAGCAAGAACAGAGCAGCACAGAAATCCCC 752
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluLys 260
Db 753 CTTCAAGCCCAATCTGATCAAGCGGCTGAGAAAGACCAAGATGAAGAGAGAAAAA 812
QY 261 GlnGluLysGluProThrLysSerProGluSerProSerProValAenSerGluThr 280
Db 813 CAAGAGAAAGAGCCCACTCCCAAGTCCCAAGATCCCGAGCAGCCAGTCAACAGTGAACA 872
QY 281 ThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSer 300
Db 873 ACATCTCTCTTCAAGAAGTTCTTCACTCACGGTTGGCGCGCTGGCGCAAGAACCCAGC 932
QY 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
Db 933 TTCAGAAATCAAAAGAGGATGATCTGGAACCTGCCAGAAAGAAAGAGCAAGAGGCA 992
QY 321 GluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGluGlnPro 340
Db 993 GAAAAAGTAGACGAGGAGAAAGAAAGAAAGAACAGACAGCCCTCGAGAGGAGCAGGACCG 1052
QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspThrGluLysValGluLeuProLeu 360
Db 1053 GCAGAAAGACACAGACCAGGCCAGGTTGTACAGACTACAGAGAGGTGGAGCTGCCTTTG 1112
QY 361 GluAspGlnValGlyAspLeuGluAlaSerSerGluLysCysAlaProLeuAlaThr 380
Db 1113 GAAGACCAGGTTGGTACCTGGAGGATCGTCAGAGAGAGTGTGCTCTTGGCAACG 1172
QY 381 GluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
Db 1173 GAAGTGTTCATGAGAAGATGGAAGCCCAAGAAAGTTGTTCAGAGGTTCCACGTGAGC 1232
QY 401 ThrValGluLysThrGluGluGluGlnGlyGlyGlyGluAlaGluGlyGlyValVal 420
Db 1233 ACCGTGGAGAAGACAGAGGAGGAGCAGGAGGAGGAGGAGGCTGAAGGGGGCGTGGTG 1292
QY 421 ValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValPro 440
Db 1293 GTAGAAGGAACAGGAGAAATCTTGGCCCTTGAGAAACTGGCTGAGCCCAAGAGGTCCCC 1352
QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
Db 1353 CAGAAGCTGAGCGTGTGAGGAGCTGATGAAGACGACAGAGATGTGTCTCTGGAGGA 1412
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
Db 1413 GACCACACTCACTGACAGACCTAAGTCTCTGAAGAGAGAGCGTGCACCAACACCCAGAA 1472
QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500

Db 1473 GGCAITTTGTCAGTGAAGTGGAGATGCTGTCTCTCAGAAAAGAAATCAAGTACAGGAAGT 1532
QY 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
Db 1533 CCCTTCAAGAAACTCTTTCAGTAGTCTCAGGCTTAAAGAAAGCTGTCTGGGAAGACAGAA 1592
QY 521 GlyLysArgGlyGlyGlyAspGluGluProGlyGluTyroClnHisIleHisThrGlu 540
Db 1593 GGGAAACGAGGAGGTGGGGAGACGAAAGAGCCCTGGAGAAATACCAACACATTCACCCGAA 1652
QY 541 SerProGluSerAlaAspGluGlnLysGlyLysSerSerAlaSerSerProGluGluPro 560
Db 1653 TCCCCAGAGAGTCTGATGAGCAGAAAGGAGAGAGTCTTCGCTGTCCTCCCGAGGAGCCT 1712
QY 561 GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
Db 1713 GAGGACACCACTGTCTCGAGAAAGGGCGCTGGAAAGCACCCCGAGATGGGAAGCTGAG 1772
QY 581 GluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAlaSerPhe 600
Db 1773 GAAGGAACCTACTTCAGTGGAGAAAGAGAGAGAGGATCACTCCCTGGGCATCCTTC 1832
QY 601 LysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGluGlu 620
Db 1833 AAAAGATGGTGACACCCCAAGAAACCGGTCCGGAAGACCTTCTGAGAGTGACAAAGAGGAA 1892
QY 621 GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
Db 1893 GAGCTGGAGAGGTCAAGAGCGCCACTTGTCTCTCACTGATAGCACAGTGTCAAGAAATG 1952
QY 641 GlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArgArgVal 660
Db 1953 CAAGATGAAGTCAAAACTGTGTGTAGAGAAACAAAGCCAGAGAAACAAAGACGTAGGGTG 2012
QY 661 AspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArgAlaArg 680
Db 2013 GATACTTCAGTGTCTTGGGAAGCAGTGTGTGTGGATCATCCCAAGAAAGAGAGCAAGG 2072
QY 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
Db 2073 AAGGCATCTCTTCAGATGATGAGGAGGGCCCAAGGACACTGGGAGGAGACACTCAGAA 2132
QY 701 AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
Db 2133 GCAGAGGAGGCGCAGCAAGAACAAAGAACCGGAAACAGACGCTGTCTCTGCAGACCCAG 2192
QY 721 GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly 740
Db 2193 GAGCAGGACCAAGCGCAAGGAAGTTCCTCACCGAGCCAGCGGAAAGCCCTTCCGAAGGG 2252
QY 741 GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
Db 2253 GAAGTGTCTCCACTTGGAGTCAATTTAAAGATTAGTCACTCCAAGAAAAAATCCAAG 2312
QY 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
Db 2313 TCAAAACTGGAAGAGAAACCGGAAGACTCTAGTGTAGAGCAGTGTCTCACTGAGATCGAA 2372
QY 781 ProSerArgGluGluSerTrpValSerIleLysLysPheIleProArgArgLysLys 800
Db 2373 CCAGTAGAAGAAATCTTGGGTTCATTTAAGAAATTCATCCCGAGCGCGGAGAA 2432
QY 801 ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluIleAsn 820
Db 2433 AGGGCAGACGGGAAGCAAGAACCAAGCCACTGTGAAGACTCAGGGCCAGTGGAGATAAT 2492
QY 821 GluAspAspProAsnValProAlaValProLeuSerGluTyroAsnAlaValGluArg 840
Db 2493 GAGGACGACCCCTAATGTCCCGAGCGTGTGCTCTCTGAGTATAATGTCAGTGGAGAG 2552
QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyroVal 860

Db 2553 GAGAAAGATGGAAGCCAGGGGAATACGAGCTGCCACAGCTGCTGGGGGCTGTGATACGTG 2612
 QY SerGluGluLeuSerIleThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
 Db 2613 TCCGAGAGAGCTCAGTAAGACTCTGTGCTCCACACTGTGATGTCGCAAGTCAATTGAGGAGCC 2672
 QY ArgAlaValThrSerValGluGluIleArgSerProSerThrIleSerAlaSerValThrGlu 900
 Db 2673 AAGGACATGACCAAGTGTCAAGAGCGGTCTCTTGGTGAATATCCGCTTCGTAACAGAA 2732
 QY ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluIleAsp 920
 Db 2733 CCTCTTGAACACACAGCGGGGAGAGGCCATGCCACTGTTGAAGAGGTCACTGAAAAAGAC 2792
 QY IleIleAlaGluGluThrProValLeuThrGlnThrGlnThrLeuProGluGlyIleAspAlaHis 940
 Db 2793 ATCATGTGACAGAAAACTCTGTGCTCACCCAGACGTACCAAGGGGTAAAGATGCCCAT 2852
 QY AspAspMetValThrSerIleValAspPheThrSerGluAlaValThrAlaThrGluThr 960
 Db 2853 GACGACATGGTACCAAGTGAAGTGAATTCACCTCAGAAAGCTGTGACAGCCACAGAGACC 2912
 QY SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyValaGluGluThrThrAsp 980
 Db 2913 TCAGAGGCTCTCCGTAACAGAAAGTACCGAAGCATCGGGGGCCGAGAGACCAAGAC 2972
 QY MetValSerAlaValSerGluLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
 Db 2973 ATGGTGTCCGCGATTTCCAGCTGACTGACTCCCGACAGACCAAGAGAGACACCCCA 3032
 QY ValGluGluValGluSerGlyValLeuAspThrGluGluGluGluGluGluGluGluGlu 1020
 Db 3033 GTTCAGAGAGTGAAGAGTGTGTGTAGATACAGAAAGAGAGAGCCAGACGAGGCC 3092
 QY IleLeuGluAlaValAlaAspIleValIleGluGluSerGluValProAlaThrGlnThr 1040
 Db 3093 ATCTCCAAAGCCCTTGCAACAGAAAGTGAAGAGAGTCCAGAGTGTGCTGCAACCCAGACT 3152
 QY ValGluArgThrGlySerIleValaLeuGluIleValaGluGluValaGluGluIleAspSerGlu 1060
 Db 3153 GTGCAGAGAGACGGGGTCAAAAGCATGAGAGAGGTGAGAGAGTGAAGAGGAGGACTCCGAA 3212
 QY ValLeuAlaSerGluIleValaAspValMetProIleGlyProValGluGluAlaGly 1080
 Db 3213 GTGCTGGCTTCGAGAAAG 3272
 QY AlaGluHisLeuAlaGlnIleSerGluThrGluGluAlaThrProGluSerLeuGluVal 1100
 Db 3273 GCTGAGCATCTTGCAACAGAGGCTTGAGACTGAGACAGGCTACTCCAGAGAGCCCTTGAAGTT 3332
 QY ProGluValThrAlaAspValaAspHisValaAlaThrCysGluValaIleIleLeuGluGln 1120
 Db 3333 CCTGAAGTACAGGCAATGATGACATGTCGACAGCTGCAAGGTATCAAGCTCCAGCAG 3392
 QY LeuMetGluGluAlaValaAlaProGluIleSerGluThrGluThrAspSerGluThrAsp 1140
 Db 3393 CTGATGAACAGAGCGCTGAGCCCTGTGATCTCGAAACCTTGACAGAGTGAACAAT 3452
 QY GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGluGlnIleAspGluThrIle 1160
 Db 3453 GGAAGACCTCCCTTACAGATTCAGACCTGCAAGATGGGACACAGCAAGATGAACATTT 3512
 QY AspSerGluAspSerIleValaAlaValaIleArgGlnSerGluValaThrGluGluGlu 1180
 Db 3513 GACAGCCAGAGACAGTAAGCACTGACAGCTGACAGCAAGTCAAGAGTCAAGAAAGAG 3572
 QY AlaAlaThrAlaGluIleValaGluGluProSerThrLeuProAspAspValProAlaGlnGlu 1200
 Db 3573 GGGGCTACTGCTCAGAAAGAGAGCCCTTGACACTACTAATATATGTCCAGGCCAGGAA 3632
 QY GluHisGlyGluGluProGluIleArgAspValLeuGluProThrGlnGluGluLeuThrAla 1220
 Db 3633 GAACATGGGGAAGAACAGGAAGATGTTCTTGAACCTACAGCAAGAGCTTACTGCT 3692

QY AlaAlaValaProValLeuAlaIleThrGluValaGluGluGluGluGluGluValaAspThrLeu 1221
 Db 3693 GCAGCCGTGCCCTTCTTGCAAGACTGAGGTGGTCAAGAGGGGTGAAGTTCAGCTGTG 3752
 QY AspGlyIleValaValaGluGluGlnIleValaPheValHisSerGlyProAspSerGln 1260
 Db 3753 GATGAGAAAAAGTCAAAAGAAAGACAGAGAGTGTGTGTACCTCTGGACCCAAAGTCA 3812
 QY LysAlaAlaAspValThrTyraAspSerGluValMetGlyValaIleGlyCysGlnIleLys 1280
 Db 3813 AAGCTCTGATGTGACATATGACAGTGAATGATGGAGATGGCCGGGTGTCAAGAAAAAG 3872
 QY GluSerThrGluValGlnSerLeuSerLeuGluGluGluGluGluGluGluGluGluGlu 1300
 Db 3873 GAGACTACTGAAGTGAAGTCAAGTTCAGTGAAGAGAGAGAGATGAAGTGAACCTTGA 3932
 QY LysGluIleValaGluThrLysProGluGlnValSerGluGluGluGluGluGluGluGlu 1320
 Db 3933 AAGGAAAAAGGAGACAAAGCCAGAGCAAGTGAAGAAAGTGAAGAGAGAGAGAGAGAG 3992
 QY AlaProGluHisGluGlyThrTyraLysProValLeuThrLeuAspMetProSerSer 1340
 Db 3993 GCTCTGAGCATGAAGAACTTACGGAAGCCAGTCTGACACTTGACATGATGCCAGCTCA 4052
 QY GluArgGlyIleValaLeuGlySerLeuGlyGlySerProSerLeuProAspGluAspLys 1360
 Db 4053 GAGAGGGGAGAGCACTGGGAAGCTTGGAGAGAACCTTCTCTCCAGACCAACAA 4112
 QY AlaGlyCysIleGluValaGlnValaGlnSerLeuAspThrThrValThrGluAlaGlu 1380
 Db 4113 GCAGTTGCATAGAGTTCATGAATCAAAAGCTTGACACACAGTCACTCAAAACAGCAGAA 4172
 QY AlaValaGluIleValaIleGluThrValaIleSerGluThrGlyGluSerProGluCys 1400
 Db 4173 GCTGTGAAAAAGTCAATGAAGACGGTGTGATTCAGAGACAGGTGAAGTCCAGAGTGT 4232
 QY ValGlyAlaHisLeuLeuProAlaGluIleSerSerAlaThrGlyGlyHisThrThrLeu 1420
 Db 4233 GTAGGTGCACTTATTCACAGCTGAAGATCTCTGTGAACGGGTGACCTGACTCTT 4292
 QY GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440
 Db 4293 CAGCATGCAAGAGACACGCTACCTCGGGGCTTGAGTCTCAGGCAAAATCCATCCCAATC 4352
 QY IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnIleGluIleSerAla 1460
 Db 4353 ATAGTAACTCCTGCTCTGAAGACACCTTACATCTGACCTTACAGAGGAATTAAGCGCA 4412
 QY SerGlnArgGluArgSerGluGluGluIleAspLysProAspAlaGlyProAspAlaAspGly 1480
 Db 4413 TCCAGAGAGAGCGATCAG 4472
 QY LysGluSerThrAlaIleGluIleValaLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
 Db 4473 AAGGAGATGACGAAATGAAAGAGTCTCAAGGCTGAACCTGAATCTTGAACTTGAG 4532
 QY SerLysSerAsnLysIleValaLeuAsnValaIleGlnThrAlaValaAspGlnPheAlaArg 1520
 Db 4533 AGTAAGAGCAACAATGATGTGCTGAACGTCAATTCAGACAGCGGTGACCAAGTTCCAGCT 4592
 QY ThrGluThrAlaProGluThrHisAlaTyraAspSerGlnThrGlnValProAlaCysArg 1540
 Db 4593 ACAGAAACAGCCCCGAAATCATATGATGATTCAGACAGACCCAGAGTTCCTGACAGAG 4652
 QY LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
 Db 4653 CTTGACAGAGGAGAGCCCAACAGATGCTGGACAAAGAAAGATGCCAAGTGAACAC 4712
 QY ProValProGlnProArgGluAspLeuGlnValLeuThrValaLeuGluValaThrAlaGln 1580
 Db 4713 CCAAGTCCGCAAGCCCAAGAGAGACTTGAAGTCTTGCAAGTCTTGGAAGGCAAGGCTGAG 4772

QY 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
|||||
Db 4773 CTTGGAAATGCTTGGCGCGTTCAGTTGAAAGCGCGGTGTCAAAG 4820

RESULT 2

RNU41453 5236 bp mRNA linear ROD 12-JAN-1999
LOCUS Rattus norvegicus PKC binding protein and substrate mRNA, complete
DEFINITION cds.
ACCESSION U41453
VERSION U41453.1 GI:4096819
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 5236)
Chapline, C., Mousseau, B., Ramsay, K. and Jaken, S.
Identification of a major protein kinase C binding protein and
substrate in rat embryo fibroblasts: Decreased expression in
transformed cells
J. Biol. Chem. (1995) In press
2 (bases 1 to 5236)
Chapline, C., Mousseau, B., Ramsay, K. and Jaken, S.
Direct Submission
Submitted (28-NOV-1995) Susan Jaken, W Alton Jones Cell Science
Center, Old Barn Road, Lake Placid, NY 12946, USA
Location/Qualifiers

FEATURES

source

CDS

/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="72"
33..4823
/note="major PKC binding protein in REF52 cells"
/codon_start=1
/product="PKC binding protein and substrate"
/protein_id="AA003788.1"
/db_xref="GI:4096820"

/translation="MGAGSSTEQSPQAGSDTPSELVLSHGHPAAEASGAAGDPAD
ADPATKLPKNGKSSVGVAEQGVHVOEENQGEVEVDVDEVDQSESDVREKDR
VREMANSTAVDITKDGOETSEIIQIPASENNVEEMVQPAESQANDVGPKYKPF
VGFKFTVKDKKEDSTVQLLVKDKDEGGAASVAGDHQSPSVETAVGESKSESE
LKQSTQKGLTKQESSTETPLQAESDQAAEEAKDEGEKQEKPEPTKSPSPSPV
NSETSPFKFTTHGAWKRTKSPKSKEDDLLETAERKEQEAKEKVDSEKTEPPA
SEEQPAEDTDQARLADYKVELPLEDQVGLDLEASSEKCAPLATEVDFDEKWEAHL
VVAEYHVSTVTEEBEQGGGAEGGVVVEGTGSELPPEKLAEPORVPOEARPAELM
KREMCVSGDHTQDLSPEKTLPHPEGIVSEVEMLSQERIKVOGSPKLKLFSS
SLKLLSKKKQKRGCGGDEEPGEYQHHTSPESADQKQESSASSPEEETTCCL
EKGPLAPQDGAEEGTSDGKKREGITPWASFQKQVTPKKRVRRPSSDKKEELEK
VKSATLSDTSVSEMDVKTVEEQKPEEPKRVDTSVSWEALICVGSKKRKA
SSSDRGGRPTLGGSHRAEASKDEAGTDVAPASTQDQQAQSSSPPEPAGSPSEG
EGSVTWSFKRLVTPRKSKSLKEKADSSVEQLSTETEPSRESWVSIKKFIPGR
KRADGKQSOATVDSGVPVIEDDPVAVVPLSEYDAVEREKVQAQNAELPQLLG
AYVSEBELKTLHVSVNAVIGTRVATVSEERSPSWISASVTEPLEHTAGAMPVE
EVTEKDIAEETPLVLTPEGKDAHDMVTSEVDTSEAVTATSETSEALTEEVTEA
SSEQETDMVSVQSLTSDPTTEATPVQVESEGVLDTEEBERTQAILQAVADVKV
SGAEVTPATQVORTQSKALEKEVEEDSEVLASEKEKQVMPKVPQVQAGAHLAQGS
ETQATPESLVEPVDVHATQVVIKLOLMEOAVAPESSETLTDSETNGSTPLA
DSDTADGTQDDETIDQDSKATAVRSQVSTEEAATAQKEPSTLPNNVPAQBEHGE
EGRDVLFTQELTAAAVFLAKTEVGEQVMDLDGKVEEKEVEFVHSGPNQKA
ADVTDSYVMGVAGQKESSTEQSLSEEGEMETDVEKEKRETPPEQVSEBGETA
APEHEGTQYKVLVTDMPSEKRGKSLGSGPSLPDQKAGCIEVQVSLDTVTOT
ABAVEKVIETVSETGESPECVGAHLPAEKSSATGGHWTLOHEDTVPLGPESQAE
SIPIIVTPAETSLHPDLOGETSASQRESSEEDKPDAGPDAGKESTAEIKVLAEP
EILEESKNKVLNVIQVAVDQFARTETAPETHAYDSQTVPMQADSEENRNCWTK
MKVAKMKHPVPOREDLQVLTLEAWAPRKCLPLRLQLKAPVSK"

BASE COUNT 1622 a 1206 c 1583 g 825 t

Alignment Scores:

Pred. No.: 9,42e-279 Length: 5236
Score: 8030.00 Matches: 1589

Percent Similarity: 99.75% Conservative: 3
Best Local Similarity: 99.56% Mismatches: 4
Query Match: 99.47% Indels: 0
DB: 10 Gaps: 0
US-09-902-432-4 (1-1596) x RNU41453 (1-5236)
QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
Db 33 ATGGGCGCAGGAGTTCACCGAGCAGCGAGCCCGGAGCAGCGCGGAGGAGCAGCAG 92
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
Db 93 CCAGGCGAGTGTGTCTAGTGGCCATGGCCGAGCTGGAAGCCCTCGGAGCAGCTGGA 152
QY 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
Db 153 GACCCCGCGGCGCGACCCGCCCAAGCTCCACAGAGAAATGCCAGCTGTCTTCT 212
QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsnGlnGluGlyGln 80
Db 213 GTCACGCGGTAGCTGAACAAGGAGATGTCCATGTCCAAGAGGAAACCCAGGAGGCGCAG 272
QY 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
Db 273 GAGGAAGAAGTCTGTGATGAGGATGTTGGACAGCGAGAGTCAAGAGATGTGAGAGAAAA 332
QY 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
Db 333 GACCGAGTGAAGAAATGGCGGCAACTCCACAGCTGTTGAAGATATCACAAGGATGGG 332
QY 121 GlnGluGluThrSerGluIleLeuGluGlnIleProAlaSerGluAsnValGluGlu 140
Db 393 CAGGAGGAGACATCAGAAATAATTGAACAGATCCCTGCTTCAGAAACAATGTGGAAGAA 452
QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
Db 453 ATGGTACAGCCCTGCTGAGTCCCAGGCTAATGATGTTGGCTTCAAGAAAGTATTAAATTT 512
QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
Db 513 GTTGGTTTAAATTCAGCTGAAGAGGATTAATAATGAAGTCAAGTACTGTCCAACTA 572
QY 181 LeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
Db 573 CTCACTGTCAAGAAGGATGAAGCGAAGGGCAGAAAGCCTCTGTGCGAGCTGGAGACCAC 632
QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLys 220
Db 633 CAGGAGCCAGTGTGGAGACTCCGTCGAGAGTCAAGCATCCAAAGAAAGTGAAGTGAAG 692
QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIlePro 240
Db 693 CAATCCACAGAGAGCAGCAAGCACCCCTGAAGCAAGAAACAGACGACGACAGAAATCCCC 752
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGluLys 260
Db 753 CTTCAAGCCGAATCTGATCAAGCGCTGAGGAAGAACCAAGATGAAGGAGAGAAAA 812
QY 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThr 280
Db 813 CAAGAAAGAGGCCCAACCAAGTCCCCCAGAAATCCCCGAGCAGGCCAGTCAACAGTGAGACA 872
QY 281 ThrSerSerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSer 300
Db 873 ACATCTCTCTTCAAGAAGTCTTCTACTACGGTGGGCCGCTGGCGCAGGAAGACGACG 932
QY 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
Db 933 TTCAGAAATCAAAAGAGAGATGATCTGGAACTGCCGAGAGAGAAAGAGAGAGGCA 992
QY 321 GluLysValAspGluGluLysGluLysThrGluProAlaSerGluGluGlnGluPro 340
|||||

Db 993 GAAAAAGTAGACGAGAGAAAAAGAGAGAGCCGCTCGGAGAGAGAGAGCCG 1052
Qy A1agiuaepThrAspGlnA1aargLeuSerA1aaspTyrgLulysValG1uleuProLeu 360
Db 1093 GCAGAGAGACACAGACAGAGCCAGGTTGTTCAGACAGACTAGAGAGGTGAGCTGCTTGG 1112
Qy 361 GluaepGlnValG1yAspLeuGlnA1aSerSerGlnGlnLysCysAlaProLeuA1aThr 380
Db 1113 GAAGACCGAGTTGTGACCTGGAGGATCTGTCAAGAGAGAAAGTGTCTCTTGGCAAG 1172
Qy 381 GluValA1PheAspGlnLysMetGlnA1aHisGlnGlnValA1aG1uValA1HisValSer 400
Db 1173 GAAGGTTTGTATGAGAGAGTGAAGCCCAAGAAAGTTGTTCAGAGAGTCCAAGTGAAGC 1232
Qy 401 ThrValG1ulysThrG1uGlnGlnGlnGlnG1yG1yG1yG1uA1aG1uG1yValVal 420
Db 1233 ACCGTGGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGAAGGGGCGTGTG 1292
Qy 421 ValG1uG1yThrG1yG1uSerLeuProG1uLysLeuA1aG1uProGlnGlnValPro 440
Db 1293 GTAGAGAGGAACGAGAGAAATCTTGGCCCTCGAGAACTGGCTGAGCCCAAGAGAGTCCCC 1352
Qy 441 GlnGlnA1aGlnProA1aGlnGlnGlnMetLysSerArgGlnMetCysValSerG1y 460
Db 1353 CAGAGAGCTGAGCTGTCTGAGAGCTGATGAAGAGCAGAGAGATGTGTCTCTGAGAGA 1412
Qy 461 AspHisThrGlnLeuThrAspLeuSerProGlnGlnLysThrLeuProLysHisProGln 480
Db 1413 GACCAACACTCACTGACACACCTTAAGTCTTGAAGAGAAACGCTGCCAAACACCCAGAA 1472
Qy 481 G1y11LevalSerGlnValA1aG1uMetLeuSerSerGlnGlnLysG1yValGlnG1ySer 500
Db 1473 GGCATTGTCTAGTGAAGTGAAGATGCTGTCTCTCAGAGAAAGATCAAGGTACAGGGAAAT 1532
Qy 501 ProLeuLysLeuLeuPheSerSerSerG1yLeuLysLeuSerG1yLysLysGlnLys 520
Db 1533 CCTTGAAGAACTCTTCTGATGACTGAGGCTTAAAGAACTGTCTGGAGAAAGCAGAAAG 1592
Qy 521 G1yLysArgG1yG1yG1yAspGlnGlnProG1yG1uTyrgLinh1s1Leh1sthrGln 540
Db 1593 GGGAAACGAGAGAGTGGGGAGAGCAGAAAGCTGGAGAAATCAACACATTCACACCGAA 1652
Qy 541 SerProGlnSerLysAspGlnGlnLysG1yG1uSerSerLysSerProGlnGlnPro 560
Db 1653 TCCCCAGAGAGCTGTGATGAGCAGAAAGAGAGAGCTGTGCTGCTCCCGAGAGGCTT 1712
Qy 561 G1uGlnThrCysLeuGlnLysG1yProLeuGlnA1aProGlnAspG1yGlnA1aGln 580
Db 1713 GAGAGAGACAGCTGTGAGAGAGGGCCGCTGGAGACCCCAAGATGGGAGAGCTGAG 1772
Qy 581 G1uG1yThrThrSerAspG1yG1uLysLysArgGlnG1y11eThrProTrpA1aSerPhe 600
Db 1773 GAAGAACTACTTCGATGGAGAGAGAGAGAGAGGGATCACTCCGCGGAGATCTTTC 1832
Qy 601 LysLysMetValThrProLysLysArgValArgArgProSerG1uSerAspLysGlnGln 620
Db 1833 AAAAAGATGTGACACCCCAAGAAACGGGTCCGAAAGACTTCTGAGAGTGCACAAAGGAGAA 1892
Qy 621 G1uLeuGlnLysValLysSerA1aThrLeuSerSerThrAspSerThrValSerGlnMet 640
Db 1893 GAGCTGGAGAGAGTCAAGAGCCGCTGTCTCTCACTGATAGCAAGTGTTCAGAAATG 1952
Qy 641 GlnAspG1yValLysThrValG1yGlnGlnGlnLysProGlnGlnLysProLysArgArgVal 660
Db 1953 CAAGATGAAGTCAAACTGTGTGTGAGAAACAAAGCCAGAGAGACCAAGAGCTTAGGGTG 2012
Qy 661 AspThrSerValSerTrpGlnA1aLeu11eCysValG1ySerSerLysLysArgA1aArg 680
Db 2013 GATACTTCACTGTCTGGAGAGCACTGATTTGTGTGATCATCTCAAGAGAGAGAGCAAG 2072
Qy 681 LysAlaSerSerSerAspAspGlnG1yLysProArgThrLeuGlnG1yLysSerHisArg 700
Db 2073 AAGGCATCTCTTCAATGATGAAGAGGGCCAAAGCACTGGAGAGGGGCACTCAACAGA 2132

Qy 701 A1aG1uGlnA1aSerLysAspLysGlnA1aG1yThrAspA1aValProA1aSerThrGln 720
Db 2133 GCAGAGAGAGCCACCAAGAGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2192
Qy 721 GluGlnAspGlnA1aGlnG1ySerSerSerProGlnProA1aG1ySerProSerGlnG1y 740
Db 2193 GAGCAGAGACCAAGGCGAG 2252
Qy 741 GlnGlnValSerThrTrpGlnSerPheLysArgLeuValThrProArgLysLysSerLys 760
Db 2253 GAAGGTGTCTCACTTGGAGAGTCAATTAAAGATTAGTCACTCCAGAGAGAGAGAGAGAGAG 2312
Qy 761 SerLysLeuGlnGlnLysValA1aGlnAspSerSerValGlnGlnLeuSerThrGln11eGln 780
Db 2313 TCMAAATCGAAG 2372
Qy 781 ProSerArgGlnGlnSerTrpValSer11eLysLysPhe11eProGlnArgArgLysLys 800
Db 2373 CCGAGTACAG 2432
Qy 801 ArgA1aAspG1yLysGlnGlnGlnA1aThrValGlnAspSerG1yProValGln11eAsn 820
Db 2433 AGGCGAGATGGAG 2492
Qy 821 GluAspAspProAsnValProA1aValValProLeuSerG1yTrpAsnA1aValGlnArg 840
Db 2493 GAGGAG 2552
Qy 841 GlnLysMetGlnA1aGlnG1yAsnThrGlnLeuProGlnLeuLeuGlnValA1aTyrgVal 860
Db 2553 GAGAGAGATGAG 2612
Qy 861 SerGlnGlnLeuSerLysThrLeuValHisThrValSerValA1aVal11eAspG1yThr 880
Db 2613 TCCAGAGAGCTCAGTAAAGACTGTGCTCACACTGAGAGTGTGCGAGTATGATGATGAGAGAG 2672
Qy 881 ArgA1aValThrSerValGlnGlnLysSerProSerTrp11eSerA1aSerValThrGln 900
Db 2673 AGGCGAGTACCAAGTGTGAAG 2732
Qy 901 ProLeuGlnHisThrA1aG1yGlnA1aMetProProValGlnGlnValThrG1uLysAsp 920
Db 2733 CCTCTTGAACACAGCCGGAG 2792
Qy 921 11e11eA1aGlnGlnThrProValLeuThrGlnThrLeuProGlnG1yLysAspA1aHis 940
Db 2793 ATCATTTGAG 2852
Qy 941 AspAspMetValThrSerGlnValAspPheThrSerGlnA1aValThrA1aThrGlnThr 960
Db 2853 GACGACATGTGTACCAAG 2912
Qy 961 SerGlnA1aLeuArgThrGlnGlnValThrGlnA1aSerG1yA1aGlnGlnThrThrAsp 980
Db 2913 TCAGAGGCTCTCGTGAAGAGATTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2972
Qy 981 MetValSerA1aValSerGlnLeuThrAspSerProAspThrThrGlnGlnA1aThrPro 1000
Db 2973 ATGCTGTCCGAGATTTCAGCTGACTGACTCCAGAGACCAAGAGAGAGAGAGAGAGAG 3032
Qy 1001 ValGlnGlnValGlnSerG1yValLeuAspThrGlnGlnGlnGlnGlnGlnGlnGlnGln 1020
Db 3033 GTTCAGAGAGTGAAGAGAGTGTCTTAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3092
Qy 1021 11eLeuGlnA1aValA1aAspLysValLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1040
Db 3093 ATCTCTCAAGCCGTTGACAG 3152
Qy 1041 ValGlnArgThrG1ySerLysA1aLeuGlnLysValGlnGlnValGlnGlnLysAspSerGln 1060
Db 3153 GTGCAAG 3212

QY 1061 ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
 DB 3213 GTGCTGCTCGGAGAAGAGAGAGAGCTTATGCGGAAAGACCCTGTCAGGAAGCTGGA 3272
 QY 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
 DB 3273 GCTGAGCATCTTGACAGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTT 3332
 QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValLleLysLeuGlnGln 1120
 DB 3333 CCTGAAGTCAGGCGAGATGAGACCATGTCGCCACGTCGCCAGGTTATCAAGCTCCAGCAG 3392
 QY 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
 DB 3393 CTGATGGAACAGGCGGTGGCCCTCGAGTCATCCGAAACCTTTCAGACAGAGTGAGACAAAT 3452
 QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
 DB 3453 GGAAGCACTCCCTTAGCAGATTACAGACTGCAGATGGGACACAGCAAGATGAACCATTT 3512
 QY 1161 AspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
 DB 3513 GACAGCCAGGACAGTAAGCCACTGTCAGCTGTCAGGAGTCACAGGTCACAGAAAGAG 3572
 QY 1181 AlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnValProAlaGlnGlu 1200
 DB 3573 CGGCTACTGCTCAGAAGAGGAGCCCTTCGACACTACTAATAATGTTCCAGCCCGCAGAA 3632
 QY 1201 GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAla 1220
 DB 3633 GAACATGGGAAGAACACAGGAAGAGATGTTCTTGAACTACACAGCAAGAGCTTACTGCT 3692
 QY 1221 AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeu 1240
 DB 3693 GCAGCCGTGCGCGTCTCGGCAAGACTGAGTGGGTCAAGAGGGTGAGGTGACTGCTGTG 3752
 QY 1241 AspGlyLysValLysGluGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
 DB 3753 GATGGAGAAAGTCAAGAAAGAACAGAGAGTGTTCATCACTCTGACCCCAACAGTCAA 3812
 QY 1261 LysAlaAlaAspValThrTyroAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
 DB 3813 AAGGCTCTGATGTGACATATGACAGTGAAGTGAAGTGGAGTGGCGGTGTGAGGAAAG 3872
 QY 1281 GluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGlu 1300
 DB 3873 GAGAGTACTGAAGTGCAGAGTCTTAGCCTCGAGAGGAGAGATGGAAACTGACGTTGAA 3932
 QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAla 1320
 DB 3933 AAGGAGAAAGGGGAGACAAAGCCAGACAGCAAGTGAAGAGAGGTGAGCAGGAACAGCC 3992
 QY 1321 AlaProGluHisGluGlyThrTyroGlyLysProValLeuThrLeuAspMetProSerSer 1340
 DB 3993 GCTCTGAGCATGAGAGAACCTACGGGAGCCAGTCTCTGACACTTTCATGATGCCAGCTCA 4052
 QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyCysProSerLeuProAspGlnAspLys 1360
 DB 4053 GAGAGGGGAGGACCTGGGAAGCCTTGGAGGAAGCCTCTCTCTCCAGACCAAGACAA 4112
 QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGlu 1380
 DB 4113 GCAGGTTGCTAGAGGTTCAAGTTCAAGGCTGGACACACAGTCACTCAACACAGCAGAA 4172
 QY 1381 AlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGluCys 1400
 DB 4173 GCTGTGAAAAGTTCATGAAAACGGTTGTGATTTTCAGAGACAGGTGAAGTCCAGAGTGT 4232
 QY 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeu 1420
 DB 4233 GTAGGTGCACACTATTATACAGCTGAGAGTCTCTCTGCAACGGGTGGCCACTGGACTCTT 4292
 QY 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle 1440

DB 4293 CAGCATGCAGAGGACACGGTACCCTGGGCGCTGAGTCTCAGCAGAGANTCCATCCCAATC 4352
 QY 1441 IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAla 1460
 DB 4353 ATAGTAACTCCTGCTCTCTGAAAGCACCCCTACATCTCTGACCTACAAGGAGAAATAAGCGCA 4412
 QY 1461 SerGlnArgGluArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGly 1480
 DB 4413 TCCAGAGAGAGCGATCAGAGGAGAGGACCAAGCCAGATGCTGCTCTGATGCTGACGGC 4472
 QY 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
 DB 4473 AAGAGAGATACACAATTCGAAAAGTCTCAAGGCTGAACCTGAGATCTCTGGAACTTGAG 4532
 QY 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
 DB 4533 AGTAAGAGCAACAAGATTGTGCTGAACGTCATTCAGACACGCCGTTCACCAAGTTCGACGCT 4592
 QY 1521 ThrGluThrAlaProGluThrHisAlaTyroAspSerGlnThrGlnValProAlaCysArg 1540
 DB 4593 ACAGAAACAGCCCCGAAACTCATGCTTATGATTCACAGACCAGGTTCTCGCAATGCAG 4652
 QY 1541 LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
 DB 4653 GCTGACAGAGGAGGCCAACAGATGCTGGACAAAATGAAGTTGCCAAGATGAACAC 4712
 QY 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGln 1580
 DB 4713 CCAGTGGCGCAGCCAGAGAGGACTTGCAGATCTCGACCGTCTTGAGGAGCATGGCTCAG 4772
 QY 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
 DB 4773 CCTCGAAATGCTTGGCGGCTTGCAGTTGAAAGCGCGGTGTCAAAG 4820
 RESULT 3
 AC115647 178954 bp DNA linear HTG 13-JUL-2002
 LOCUS Rattus norvegicus clone CH230-171G16, *** SEQUENCING IN PROGRESS
 DEFINITION *** 53 unordered pieces.
 ACCESSION AC115647
 VERSION AC115647.3 GI:21737529
 KEYWORDS HTG; HTGS PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 178954)
 AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
 Barbarta, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Carrott, F.F., Carter, M., Cavazos, S.K., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
 Homs, J., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louseged, H.,
 Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,

```

Massey, E., Mawhiney, E., McLeod, M. P., Meador, M., Mei, G., Meeker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogun, M., Okunolu, G.,
Orangun, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherrer, S., Scott, G., Shen, H., Shooshkat, N., Sisson, I.,
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umani, S., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, K., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Unpublished
2 (bases 1 to 178954)
Worley, K. C.
Direct Submission
Submitted (22-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178954)
Worley, K. C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:21239989.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GPHP
Center clone name: CH230-171G16
----- Summary Statistics
Sequencing vector: plasmid;
Assembly program: Phrap, version 0.990329
Consensus quality: 142023 bases at least Q40
Consensus quality: 146696 bases at least Q30
Consensus quality: 150215 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1459: contig of 1459 bp in length
1460
1559: gap of unknown length
1560
2931: contig of 1372 bp in length
2932
3031: gap of unknown length
3032
4420: contig of 1389 bp in length
4421
4520: gap of unknown length
4521
5800: contig of 1280 bp in length
5801
5900: gap of unknown length
5901
7559: contig of 1659 bp in length
7560
7659: gap of unknown length
7660
8887: contig of 1228 bp in length
8888
8987: gap of unknown length
8988
10297: contig of 1310 bp in length
10298
10397: gap of unknown length
10398
12090: contig of 1693 bp in length
12091
12190: gap of unknown length

12191
13523
13623
15623
15630
17541
17541
17641
19820
19820
19920
21261
21262
21362
23786
23786
23886
25653
25653
25653
27443
27443
27543
30416
30516
31785
31785
31885
34971
35071
36234
36234
36334
39669
39669
39769
42867
46088
46088
46189
46189
49028
49028
49128
51580
51680
55193
55292
55292
55293
58205
58304
60939
61039
61040
6185
64285
64285
67777
67777
67877
72019
72020
72120
72147
75148
75148
75248
78406
78406
82238
82238
82339
82339
85154
85154
87767
87767
87768
87768
87868
90643
90643
90743
90743
94014
94014
94113
94113
98294
98294
98394
98394
101682
101682
101782
101782
101783
104237
104237
104337
104337
108706
108706
108806
108806
113208
113208
113208
113209
113209
117239
117239
117339
117339

13522: contig of 1332 bp in length
13622: gap of unknown length
15623: contig of 2007 bp in length
15629: gap of unknown length
17540: contig of 1811 bp in length
17540: gap of unknown length
19819: contig of 2179 bp in length
19819: gap of unknown length
19919: gap of unknown length
21261: contig of 1342 bp in length
21261: gap of unknown length
21361: gap of unknown length
23785: contig of 2424 bp in length
23785: gap of unknown length
25662: contig of 1677 bp in length
25662: gap of unknown length
25662: gap of unknown length
27442: contig of 1780 bp in length
27442: gap of unknown length
27542: gap of unknown length
30415: contig of 2873 bp in length
30415: gap of unknown length
30515: gap of unknown length
31784: contig of 1269 bp in length
31784: gap of unknown length
31884: gap of unknown length
31970: contig of 3086 bp in length
31970: gap of unknown length
35070: gap of unknown length
36233: contig of 1163 bp in length
36233: gap of unknown length
36333: gap of unknown length
39668: contig of 3335 bp in length
39668: gap of unknown length
39768: gap of unknown length
42866: contig of 3098 bp in length
42866: gap of unknown length
46088: contig of 3122 bp in length
46088: gap of unknown length
49027: contig of 2839 bp in length
49027: gap of unknown length
49127: gap of unknown length
51579: contig of 2452 bp in length
51579: gap of unknown length
51679: gap of unknown length
55192: contig of 3513 bp in length
55192: gap of unknown length
55292: gap of unknown length
58204: contig of 2912 bp in length
58204: gap of unknown length
58304: gap of unknown length
60939: contig of 2635 bp in length
61039: gap of unknown length
6184: contig of 3145 bp in length
6184: gap of unknown length
64284: gap of unknown length
67776: contig of 3492 bp in length
67776: gap of unknown length
67876: gap of unknown length
72019: contig of 4143 bp in length
72019: gap of unknown length
75147: gap of unknown length
75147: contig of 3028 bp in length
75247: gap of unknown length
78405: contig of 3158 bp in length
78405: gap of unknown length
78505: gap of unknown length
82238: contig of 3733 bp in length
82238: gap of unknown length
82338: gap of unknown length
85153: contig of 2815 bp in length
85153: gap of unknown length
85253: gap of unknown length
87767: contig of 2514 bp in length
87767: gap of unknown length
87867: gap of unknown length
90643: contig of 2776 bp in length
90643: gap of unknown length
90743: gap of unknown length
94013: contig of 3270 bp in length
94013: gap of unknown length
94113: gap of unknown length
98293: contig of 4180 bp in length
98293: gap of unknown length
98393: gap of unknown length
101682: contig of 3289 bp in length
101682: gap of unknown length
101782: gap of unknown length
104237: contig of 2455 bp in length
104237: gap of unknown length
104337: gap of unknown length
108706: contig of 4369 bp in length
108706: gap of unknown length
108806: gap of unknown length
113208: contig of 4402 bp in length
113208: gap of unknown length
117238: contig of 3930 bp in length
117238: gap of unknown length
122577: contig of 5239 bp in length

```

* 122578 122677: gap of unknown length
 * 122678 126779: contig of 4102 bp in length
 * 126780 126779: gap of unknown length
 * 126880 131807: contig of 4928 bp in length
 * 131808 131907: gap of unknown length
 * 131908 138315: contig of 6408 bp in length
 * 138316 138415: gap of unknown length
 * 138416 143597: contig of 5182 bp in length
 * 143598 143697: gap of unknown length
 * 143698 152755: contig of 9058 bp in length
 * 152756 152855: gap of unknown length
 * 152856 157897: contig of 5042 bp in length
 * 157898 157997: gap of unknown length
 * 157998 167260: contig of 9263 bp in length
 * 167261 167360: gap of unknown length
 * 167361 178954: contig of 11594 bp in length.

FEATURES
 Location/Qualifiers
 source 1..178954

Alignment Scores:

Pred. No.: 4,88e-257 Length: 178954
 Score: 7461.00 Matches: 1488
 Percent Similarity: 99.20% Conservatives: 6
 Best Local Similarity: 98.80% Mismatches: 11
 Query Match: 92.42% Indels: 3
 DB: 2 Gaps: 0

US-09-902-432-4 (1-1596) x AC115647 (1-178954)

QY 89 ValGlyGlnArgGluSerGluAspValArgGluLysAspArgValGluGluMetAlaAla 108
 Db 174384 GTTGGACGGCAGAGTCTGAGAGATGTCAGAGAAAGAACCGAGTGAAGAAATGGCGGCC 174443
 QY 109 AsnSerThrAlaValGluAspIleThrLysAspGlyGlnGluGluThrSerGluIle 128
 Db 174444 AACTCCACAGCTGTTGAAGATATCACAAAGATGGGACAGAGGAGACATCAGAAATATT 174503
 QY 129 GluGlnIleProAlaSerGluAsnAsnValGluGluMetValGlnProAlaGluSerGln 148
 Db 174504 GAACAGATCCCTGCTTCAGAAAGCAATGTGTGAAGAAATGGTACAGCCCTGCTGATCCAG 174563
 QY 149 AlaAsnAspValGlyPheLysLysValPheLysPheValGlyPheLysPheThrValLys 168
 Db 174564 GCTAATGATGTGGCTTCAGAAAGATATTTAAATTTGTTGTTTAAATTCACGGTGAAG 174623
 QY 169 LysAspLysAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGly 188
 Db 174624 AAGGATAAAATGAAGATCAGATACCTGTCNACTACTCTACTGTCAAGAGGATGAAGGC 174683
 QY 189 GluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGluThrAla 208
 Db 174684 GAAGGGGCAGAAGCCTCTGTCGGAGCTGGAGACCACCCAGGAGCCAGTGTGGAGACTGCC 174743
 QY 209 ValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGly 228
 Db 174744 GTCGGAGAGTCAGCATCCAAAGAAAGTGAGCTGAAGCAATCCACAGAGAAGCAAGAGGC 174803
 QY 229 ThrLeuLysGlnGluGlnSerSerThrGluIleProLeuGlnAlaGluSerAspGlnAla 248
 Db 174804 ACCCTGAAGCAGACAGACAGCAGCAGAAATCCCCCTTCAAGCCGATCTGTATCAAGCG 174863
 QY 249 AlaGluGluAlaLysAspGluGlyGluGluLysGlnGluLysGluProThrLysSer 268
 Db 174864 GCTCAGGAAGAAGCAAGATGAAGAGAGAGAAACCAAGAGAAAGGCCACCAAGTCC 174923
 QY 269 ProGluSerProSerSerProValAsnSerGluThrThrSerSerPheLysPhePhe 288
 Db 174924 CCAGAAATCCCCGAGACGCCAGTCAAGAGTGAAGCAACATCTCTTCAAGAAATGTTCTTC 174983
 QY 289 ThrHisGlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspAsp 308
 Db 174984 ACTCAGGTTGGGCGGCTGGCGCAGAGACCCAGCTTCAAGAAATCAAAGAGAGATGAT 175043

QY 309 LeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluGluGluLys 328
 Db 175044 CTGGAACCTGCCAGAGAGAGAAAGAGGAGCAAGAGCGAGAAAGTAGACGAGGAGAAAG 175103
 QY 329 GluLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArg 348
 Db 175104 GAAAGACAGAGAGCAGCCTCGAGAGGAGAGAGCCGCGAGAGACACAGACACAGCCAGCCAGG 175163
 QY 349 LeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
 Db 175164 TTGTCAGCAGACTACGAGAGAGTGGAGCTGCCCTTGGAGAGCAGGTTGCTGACCTGGAG 175223
 QY 369 AlaSerSerGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
 Db 175224 GCATCGTCAGAGAGAGAGTGTGCTCTCTTTGGCAACGGAAGTGTTCATGAGAAAGATGGAA 175283
 QY 389 AlaHisGlnGluValValAlaGluValHisValSerThrValGluLysThrGluGluGlu 408
 Db 175284 GCCCACCAAGAGTTGTTGAGAGAGTCCAGCTGAGCAGCCGTTGGAGAGACAGAGGAGGAG 175343
 QY 409 GlnGlyGlyGlyGlyGluAlaGluGlyGlyValValValValGluGlyThrGlyGluSerLeu 428
 Db 175344 CAGGAGGAGGAGAGAGAGCTGAAGGGGGCGTGTGTGTAGAGAGACAGAGAGATCCTTG 175403
 QY 429 ProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
 Db 175404 CCCCCTGAGAAACTGGCTGAGCCCGCAGAGGTCCTCCCGAGAAAGCTGAGCCTGCTGAGGAG 175463
 QY 449 LeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeu 468
 Db 175464 CTGATGAAGAGCAGAGAGATGTGTGTCCTGGAGAGAGACACACTCAACTGACAGACCTA 175523
 QY 469 SerProGluGluLysThrLeuProLysHisProGluGlyLysValSerGluValGluMet 488
 Db 175524 AGTCCTGAGAGAGAGAGCTGCCCCAACACCCAGAGAGCATTTGTCAGTGAGGTGGAGATG 175583
 QY 489 LeuSerSerGlnGluArgLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
 Db 175584 CTGTCCTCTCAGGAAAGATCAAGGTACAGGGAAGTCCCTTGAAGAAACTCTTTCAGTAGC 175643
 QY 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysAsp 528
 Db 175644 TCAGGCTTTAAAGAGCTGTCTGGAGAGAGAGAGAGGGGAAACGAGGAGGTGGGGAGAC 175703
 QY 529 GluGluProGlyGluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGln 548
 Db 175704 GAAGACCTGGAAGATACCAACACATTCACACCGAATCCCCAGAGAGTCTGTATGAGCAG 175763
 QY 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLys 568
 Db 175764 AAGGGAGAGAGCTCTGCTGCTGCTCCCGAGGAGCCTGAGGAGACCACTGCTCTGGAGAAA 175823
 QY 569 GlyProLeuGluAlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGlu 588
 Db 175824 GGCCCTCGAAGACACCCAGGATGGGGAAGCTGAGAGAGAGAGCTGAGAGAGGCTCAAGAGCGCC 175883
 QY 589 LysLysArgGluGlyLysThrProTyrAlaSerPheLysLysMetValThrProLysLys 608
 Db 175884 AGAAGAGAGGAGAGATCACTCCCTGGGATCCCTTCAAAAGATGTTGACACCCCAAGAAA 175943
 QY 609 ArgValArgArgProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAla 628
 Db 175944 CGGGTCCGAAGACCTTCTCAGAGATGACAGGAGAGAGAGCTGAGAGAGGCTCAAGAGCGCC 176003
 QY 629 ThrLeuSerSerThrAspSerThrValSerGluMetGluAspGluValLysThrValGly 648
 Db 176004 ACCTTCTCTCCACTGATAGCAGTGTGAGAAATGCAAGATGAAGTCAAACTGTTGTT 176063
 QY 649 GluGluGlnLysProGluGluProLysArgValAspValAspThrSerValSerTrpGluAla 668
 Db 176064 GAGGAACAAAGCCAGAGAGAACCAAGCCGTAGGGTGGATACCTTCAGTGTCTTGGGAAGCA 176123
 QY 669 LeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGlu 688

Dh	176124	CTGATTGTGTGCGATATCCAGAGAGACCAAGAGGCGATCTTCCATGATGAA	176183
Qy	689	GIYGIYProArGThrLeuGIYGIYAspSerHisArgAlaGluGluAlaSerLysAspLys	708
Dh	176184	GGAGGGCCAGAGACCTGGAGGGGACGATCACAGAGCAGAGGGCCAGCAAAAGACAA	176243
Qy	709	GIYAlaGIYThrAspAlaValProAlaSerThrgInGluInaSerGlnGIYSer	728
Dh	176244	GAAGCCGGAACAGACGCTGTCTCTGCGACCCAGAGAGCGACCAAGCGCAAGGAAGT	176303
Qy	729	SerSerProGluInProAlaGIYSerProSerGluGIYGIYValSerThrTrpGluSer	748
Dh	176304	TCCCTACCCGAGCCAGCGGGAAGCCCTTCCAGAGGGAGGTGTCTCCACTTGGGAATCA	176363
Qy	749	PhelYSArgLeuValThrProArgLysSerLysSerLysLeuGluGluValAlaGlu	768
Dh	176384	TTTAAAGAATTTAGTACTCCAGAAAAAAATCCAAAGTCAAACCTGGAGAGAAAGCCGAA	176422
Qy	769	AspSerSerValGluGluInLeuSerThrgInLeuProSerArgGluGluSerTrpVal	788
Dh	176424	GACTTAGTGTAGACAGAGTGTGCCACTGATGTCGACCGAGTGAAGAAATCTTGAGTT	176483
Qy	789	SerTIIelysLysPheIleProGIYArgArgLysLysArgAlaAspGIYLySGInGluIn	808
Dh	176484	TCCATTAAAGAAATTCATCTCCCGGACGCGCGAAGAAAGGCGAGATGGGAAGCAAGAA	176543
Qy	809	AlaThrValGluAspSerGIYProValGluIleAsnGluAspAspProAsnValProAla	828
Dh	176544	GCCACTGTGGAAGACTTCAGGGCCAGTGGAGATTAATGAGACGACCTTGATGTCCAGCC	176603
Qy	829	ValValProLeuSerGIYTrpAsnAlaValGluArgGluLysSmetGluAlaGInGIYAsn	848
Dh	176604	GTCGTGCTCTGTGTGAGATGATGTCAGTGGAGAGGAGAGATGGAAGCCAGGGGAAT	176663
Qy	849	ThrGluLeuProGluInLeuGluValAlaValTrpValSerGluGluLeuSerLysThrLeu	868
Dh	176664	GCGGAGCTGCCCCACACTCTGTGGGGCTGTGTACGTGTCCGAGAGCTCAGTAAAGACTCTG	176723
Qy	869	ValHisThrValSerValAlaValIleAspGIYTrpArgAlaValThrSerValGluGlu	888
Dh	176724	GTCACACTGTAGTGTGCACTCACTATGAGGACCAAGGCAAGTCAACAGTGTGAAAGAG	176783
Qy	889	ArgSerProSerTrpIleSerAlaSerValThrgInProLeuGluHisThrAlaGIYGIY	908
Dh	176784	CGGTCTCTCTGCGATATCCGCTTCGTTAACAGAACTCTTGAACACACGGCGGAGAA	176843
Qy	909	AlaMetProProValGluGluValThrgLysAspIleIleAlaGluGluThrProVal	928
Dh	176844	GCCATGCCACCTGTTGAAGAGGTCACTGAAAAAGCATCATTTGCAAGAAATCTCTGTG	176903
Qy	929	LeuThrGluInThrLeuProGluGIYLysAspAlaHisAspAspMetValThrSerGluVal	948
Dh	176904	CTCACCCAGACCTTTCACAGAGGTAAAGATGCCATGCGACATGTGTCCAGATGAAGTG	176963
Qy	949	AspPheThrSerGIYAlaValThrAlaThrGluThrSerGIYAlaLeuArgThrgInGlu	968
Dh	176964	GATTTTCACCTCAGAAAGCTGTGACAGCCACAGAGACTTCAAGAGCTCTCCATCTGAAGA	177023
Qy	969	ValThrgInAlaSerGIYAlaGluGluThrThrAspMetValSerAlaValSerGluLeu	988
Dh	177024	GTTTACCGAAGCATCGGGGGCCGAAAGACCAACAGACATGTGTGTCCGAGTTTCCACCTG	177083
Qy	989	ThrAspSerProAspThrThrgInGluGluAlaThrProValGIYGIYValGIYSerGIYVal	1008
Dh	177084	ACTGACTCCCCAGACACCAACAGAGAACCCACCCAGTTCAGAGGTGAGAGGGTGTGTG	177143
Qy	1009	LeuAspThrGluGluGluGluArgGluInThrgInGluAlaIleLeuGluAlaValAlaAspLys	1028
Dh	177144	CTAGATTACAAAGAAAGAGAGCGCCAGACGAGGCCATCTTCCAAAGCCGTTCAGACAAAG	177203
Qy	1029	ValLysGluGluSerGluValProAlaThrGluInThrValGIYArgThrgLysSerValAla	1048

Dp	177204	GTAAAGAGAGTCCCAAGTGGCTCTGCACACCCACACTGTGTGACGAGAACGGGGCTCAAAAGCA	177203
Qy	1049	LeuGluValValGluGluValGluGluAspSerGluValLeuAlaSerGluValGluValys	1068
Dp	177264	CTGGAGAAAGTTTGAGAGAGTATGAGAGAGACTCCGAAAGTGTGGCTTCGGAGAAAGAAAG	177323
Qy	1069	AspValMetProIysSerGlyProValGlnGluValGlyValGluHisLeuAlaGlnGlySer	1088
Dp	177324	GACGTATGCGCGAAAGAACCCGTGTGCAGAGAAAGCTGAGCTGATCTTGCACAGGGCTCTC	177383
Qy	1089	GluThrGlnGlnAlaThrProGluSerLeuGluValProGluValThrAlaAspValAsp	1108
Dp	177384	CAGACTGGACAGCGCTACTCCAGAGAGCCCTTGAAAGTTCTTAAGTTCACAGCAGATGTAGAC	177443
Qy	1109	HisValAlaThrCysGlnValIleIleIleLeuGlnGlnLeuMetGlnAlaValAlaPro	1128
Dp	177444	CATGTCCGCACTGTCAGAGTTATCAAGCTCCAGAGCTGATGGAACAGGCGTGGCCCCCT	177503
Qy	1129	GluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAspSer	1148
Dp	177504	GAGTCATCCGAAACCTTGACACGACAGTGAACAATGGAAGCACTCCCTTAGCAGATTCA	177563
Qy	1149	AspThrAlaAspGlyThrGlnGlnGlnAspGluThrIleAspSerGluAspSerIleAlaThr	1168
Dp	177564	GACACTGCAGATGGAGACACGCAAGATGAACCAATTGACGCCAGACAGTAAAGCCACT	177623
Qy	1169	AlaAlaValAlaArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnValSGluGlu	1188
Dp	177624	GCAGCTGTCAAGCGATGCACAGCTGCACAGAAAGAAAGAGCGCGCTCTCTCAGAAAGAGAG	177683
Qy	1189	ProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGlyArg	1208
Dp	177684	CCCTTCGACACTCACTAATATATGTTCACGCCAGAGAAAGATGGGAGAAACCGAGAAAG	177743
Qy	1209	AspValLeuGluProThrGlnGlnGlnLeuThrAlaAlaAlaValProValLeuAlaValys	1228
Dp	177744	GATGTTCTTGAACTTACACAGCAAGCAAGACTTGTGTGTGCACCGCTGCCCCCTTCGGCAAG	177803
Qy	1229	ThrGluValGlyGlnGluGluGluValAspTrpLeuAspGlyGluValysSGluGlu	1248
Dp	177804	ACTGAGGTGGGTCAAGAGGGTGAAGGTTGACTGTGTGATGGAAGAAAGCAAAAGAAAGAA	177863
Qy	1249	GlnGluValPheValHisSerGlyProAsnSerGlnValAlaAlaAspValThrTyrAsp	1268
Dp	177864	CAGAGGTGTTGTGACACTGTGCAGCCCAACAGTCAAAAGGCTGCTGATGTGACATATGAC	177923
Qy	1269	SerGluValMetGlyValAlaGlyCysGlnGlnGluSerThrGluValGlnSerLeu	1288
Dp	177924	AGTGAAGTGAAGGAGAGTGGCCGGGTGTCAAGAAAGAGAGTCAAGAGCAAGACTCTT	177983
Qy	1289	SerLeuGlnGluGluGluMetGluThrAspValGluValGluValysArgGluThrTyrPro	1308
Dp	177984	AGCTGGAGGAGGAGAGATGGAACCTGACGTTGAAAAGAGAAAGAGAGCAAAAGCCA	178043
Qy	1309	GlnGlnValSerGluGluGluGluGlnGluThrAlaAlaProGluHisGluGlyThrTyr	1328
Dp	178044	GAGCAAGTGAAGTGAAGAGGTGAACAGAAACAGCGCGCTCTCTGAGCATGAAGAACCTAC	178103
Qy	1329	GlyIysProValLeuThrLeuAspMetProSerSerGluArgGlyIysValAlaLeuGlySer	1348
Dp	178104	GGGAAGCCAGTCTTGACACTTGACATCCCAAGCTCAAGAGGGGGAAAGGCACTGGGAAGC	178163
Qy	1349	LeuGlyIysSerProSerLeuProAspGluAspIleValGlyCysIleGluValGlnVal	1368
Dp	178164	CTTGGAGGAAGCCCTTCTCTCCACAGACCAAGAACAAACAGATGTGCATAGAGGTTCAAGTT	178223
Qy	1369	GlnSerLeuAspThrThrValThrGlnThrAlaGluValAlaGluValysValIleGluThr	1388
Dp	178224	CAAAACCTGGACACACAGCTACTCAAAACAGCAGAAACCTGTGGAAAAAGTCATAGAAACG	178283
Qy	1389	ValValIleSerGluThrGlyIysSerProGluCysValGlyValHisLeuLeuProAla	1408
Dp	178284	GTGTGATTTTCAAGACACGCTGGAAGTTCACAGAGTGTGTAGTGTGCACACTTATTAACAGCT	178343

```
QY 1409 GluLysSerSerAlaThrGlyHisTrpThrLeuGlnHisAlaGluAspThrValPro 1428
|
|
|
Db 178344 GAGAAGTCTCTCAACGGGTGGCCACTGGACTCTTCAGCATGCAGAGACACGGTACCC 178403
|
|
|
QY 1429 LeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSer 1448
|
|
|
Db 178404 CTGGGGCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAATCTCTCTCTGAAGC 178463
|
|
|
QY 1449 ThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGlu 1468
|
|
|
Db 178464 ACCCTACATCTGACCTACAGGAGAAATAGCGCATCCAGAGAGAGCGATCAGAGAA 178523
|
|
|
QY 1469 GluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLys 1488
|
|
|
Db 178524 GAGGACAAAGCCAGATGCTGCTCTGATCTGACGGCAAGGAGAGTACAGCAATCGAAAA 178583
|
|
|
QY 1489 ValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerLysIleValLeu 1508
|
|
|
Db 178584 GTCTCAAGGCTGAACCTGAGATCTCTGAACTTGAGAGTAAAGACCAAGATTGTGCTG 178643
|
|
|
QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHis 1528
|
|
|
Db 178644 AACGTATTACAGACGCGTTGACAGTTCCGACGTACAGAAACAGCCCGGAACTCAT 178703
|
|
|
QY 1529 AlaTrpAspSerGlnThrGlnValProAla-CysArgLeuAspSerArgGluProAsnAr 1548
|
|
|
Db 178704 GCTTATGATTACAGACCCAGGTTCTGCAATGCAGCTGACAGCCAGGAGGCCAACAG 178763
|
|
|
QY 1548 GYsTrpThrLysMetLysAspAlaLysMetLysHisProValProGlnProArgGluAs 1568
|
|
|
Db 178764 ATCTGGACAAAAATGAAAGTTCCCAAGATGAAACACCCAGTCCCGCAGCCAGAGAG 178823
|
|
|
QY 1568 pLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCysLeuProArgLe 1588
|
|
|
Db 178824 CTTGCAAGTCTGACCGTTCTTGAGCA-TGGGCTCAGCCTCC-GAATGCTTGGCCGGCT 178881
|
|
|
QY 1588 uGlnLeuLysAlaPro 1593
|
|
|
Db 178882 GCANNTGAAAGCGCG 178897
|
|
|
RESULT 4
AB020886 Mus musculus mRNA for SSeCKs, complete cds. linear ROD 14-APR-1999
LOCUS
DEFINITION
ACCESSION AB020886
VERSION
KEYWORDS SSeCKs.
SOURCE tissue_lib:Clontech mouse testis library cDNA to mRNA, clone:133.5.
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Okita,K., Kitamura,H., Iwanaga,T., Morimatsu,M. and Saito,M.
Isolation and characterization of mouse SSeCKs.
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 6195)
Okita,K., Kitamura,H., Saito,M. and Morimatsu,M.
Direct Submission
JOURNAL Submitted (08-DEC-1999) Keisuke Okita, School of Veterinary
Medicine, Hokkaido University, Laboratory of Biochemistry; N18, W9,
Kita-ku, Sapporo, Hokkaido 060-0818, Japan
(E-mail:v080008@vet.ec.hokudai.ac.jp, Tel.81-11-706-5206,
Fax:81-11-757-0703)
COMMENT Sequence updated (10-Mar-1999).
FEATURES
source
1. .6195
/organism="Mus musculus"
/strain="BALB/c"
/isolate="200"
/db_xref="taxon:10090"
/clone="133.5"
```

```
/sex="male"
/tissue_type="testis"
/dev_stage="9-12 weeks"
/tissue_lib="Clontech mouse testis library"
1. .6195
/gene="ssecs"
/locus="5194"
/gene="ssecs"
/note="Src Suppressed C Kinase Substrate"
/codon_start=1
/product="SseCKs"
/protein_id="BAA76894.1"
/db_xref="GI:4589832"
/translation="MGAGSSTEQRSPEQAPASDTPSELSHGHGPAABASGAGPAD
ADPATKLQKQQLSVAQVAVAEQEDVHVEESQDQEEVTVEDVQSESDVSKDKR
AKWAASSTVDEITKDRQETPEIIEQIPASESNEVMAQAASQANDVGFQKVFKE
VGFPTVKKDKNEKSTVOLLTKVDEGEAGASVAGDHOEPGVETVCEASKESEL
KOSTEKORGTLOAQSSTEIPLQAESGGTTEERAAKDGEEENREKEPTKPLESPTSPV
NETSSPFKFTTHWAGWRKTSFKPKEDDLETSKKEQEAKEVDEEGEKTBPAP
AEQEPASGTDOARLSADYKVEUPLDQVGDLEALSEKCAPLATEVEDEKTEAHOEV
VAEVHSTVEKMTKQGGAEVGVVVEGSGESLPEKLAETQVEVPAEPVEELMKT
KEVCSGCDHTQLTDLSPERKMLPKHPEGIVSEVEMLSQERIKVQGSPLKKLFSSG
LKLKSGKQKGRGGGDEBPGEYHOITQTPESADQKGESSASSPEEPEIACLE
KPSAPQEAEEAGATSDGKKEGITPMASFKMVTPKKVRPSRPSDEKEELDVKV
KSALSTESTASGMDVRAVGEORSEEPKRVDTVSWEALLICVSSSKRKARAS
SSDDGGPRTLGGDGHRAEASAKKEADALPASTQEQDQAHGSSSPSPAGSPSEGGV
STWESFKLVTPRKSKSLERAEDSGAEQLASEIPEFSREESWVSIKKFIPGRKKR
ADGQEOAAVEDSGPEINEDPDVPAVVPLSEYDAVERKLEAQAQENBELPOLKG
AVTYSEELSKTLVHTVSVAVIDGTRAVTSARESPSWISASMTPELEHAEGVATPPVG
EVTEKDITAEATPALAQLPGKQADDDIVTSEVDFTSSEAVTAAETALRAEELTEA
SGAEETDMVSQVLSQSDPTTEETATPQVEEGMLDTEQEROTQAVLQAVADVKV
EDSQPATQTLQRAQKALEKVEEEDSEVLATEKEDKVVPEGVQEAETHELAQGS
ETQATPELSLEVEPTEDRATTCCVTKHQLMQEAQVAPESSELTLTSETNGSTPLA
DSDTPNGTQDQETVDSQDNAAIAVKQSVTTEEAQAQTEGTPSPSPAPQBEHREK
PCRDVLEPTQALAGAVPILAKAEVQSGEAGQFQGEVKQGCQVKELEVPVHTGPNS
QKADLTDRDSVMEVARCQETESNEEQSISPEKRMGTVDKEKETETKTEQASBEHQ
ETAAPDEHSGTHPVLITADMPHSERGKALGSLGSPDQKADICIEVQVQSSDTPV
TQTTEAVKKVEETVATSEMSLESCAGASLPAEKLSEKTYGTYLQHGEDTVPQGPES
QAESPIITVPAPESILHSDILQREVSQKORSDNDKPDAGDPAAGKESAAAREKILR
AEPEILEKSNKIVQSVIOTAVDQFARTAPETHASDLQNVPMOADQSGAQOM
LDKDSDLQVSPQDCTLAVAGEGLAVSDSGSKASEMITTLAVESAVKESVEKL
PLOCKDEKEHAADGPHQSLAKAEADASGNLTKESTPDNPGKLTTEGDALKEMNKAQ
TEEDDLQBPKGDLTES"
6195
/gene="ssecs"
/note="20 A nucleotides"
BASE COUNT 1838 a 1451 c 1777 g 1129 t
ORIGIN
Alignment Scores:
Pred. No.: 7,31e-230 Length: 6195
Score: 6663.00 Matches: 1357
Percent Similarity: 88.90% Conservative: 76
Best Local Similarity: 84.18% Mismatches: 156
Query Match: 82.53% Indels: 23
DB: 10 Gaps: 11
US-09-902-432-4 (1-1596) x AB020886 (1-6195)
QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
|
|
|
Db 140 ATGGGTGCAGGAGTTCACCGAGCAGCGAGCCCGCAGCAGCGCGGAGCGACACG 199
|
|
|
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGly 40
|
|
|
Db 200 CCGAGCGAGCTGGAGCTTCAGTGCCCATGGCCCGCAGCGAGCGCGGAGCGAGCTGGA 259
|
|
|
QY 41 AspProAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSer 60
|
|
|
Db 260 GATCCCGCTGACGCGGACCCCGCCACCAAGCTCCACAGAGAGATGGTCAGCTGTCTGCC 319
|
|
|
QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsnGlnGluGlyGln 80
|
|
|
```


Dh 320 GTCAATGGTGTAGCTGAACAGAAAGATGTCCAGTCCAAAGAGAAAAGCCAGATGGGCAA 379
Qy 81 GtuglugluValaIAspGluAspValGlyGlnArgGluSerGluAspValArgGlnuys 100
Dh 380 GAGGAAAGAGTACCTGTTGAAGATGTTGACAGAGAGATCAGAAAGATGTGAAGAAAAA 439
Qy 101 AsparGValGluGluMetAlaIAsnSerThrAlaValGluAspLLeuThrLysAspGly 120
Dh 440 GACCAGCTTAAAGAAATGCGACCGAGTTCCACAGTTGTGGAAGATATCTCAAAAGGCGAG 499
Qy 121 GtnglugluThrSerGluLeuIleGluGluIleProAlaSerGluAsnAsnValGluGlu 140
Dh 500 CAGGAGGAAACCCGAAATTAATCGAAAGATCCCTGCTTCAGAGAGCAATGTGAAGAA 559
Qy 141 MetValGlnProAlaGluSerGlnAlaIAsnAspValGlyPheLysLysValPheLysPhe 160
Dh 560 ATGGCGCAGGCTGCTGATGCCAAGCTAATGACGTGCGCTTCAAGAAAGTATTTAAATT 619
Qy 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
Dh 620 GTTGGTTTAAATTCACGCGTGAAGAGATTAATAAAGCAAAAGTCAGATACCGTCCAGCTA 679
Qy 181 LeuThrValLysLysAspGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 200
Dh 680 CTCACGTGTCAAGAGATGAAGCGAAGGCGCAAGAGCCCTCGGTGAGCAGAGACAC 739
Qy 201 GtngluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
Dh 740 CAGAGCCCGAGGTGAGAGACC---GTCCGCGAATCAGATCCAAAGAAATGAGCTGAAG 796
Qy 221 GtngluThrGluLysGlnGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 240
Dh 797 CATCCACAG 856
Qy 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluGluGluGluGluGluGluGluGlu 260
Dh 857 CTTCAAGCCGAACTGTGTCAAGGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 916
Qy 261 GtngluLysGluProThrLysSerProGluSerProGluSerProValAsnSerGluThr 280
Dh 917 CGAGAGAAAG 976
Qy 281 ThrSerSerPheLysLysPhePheThrHisGlyTyrAlaGlyTyrAlaGlyLysLysThr 300
Dh 977 ACATCTTCTTCAAGAAATCTTCACTCAAGCGCTGCGGCGGCTGCGGCAAGAGAGAG 1036
Qy 301 PheLysLysSerLysGluAspLeuGluGluGluGluGluGluGluGluGluGluGluGlu 320
Dh 1037 TTCAGAAAG 1096
Qy 321 GtngluValAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 339
Dh 1097 GAAAG 1156
Qy 340 ProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuPro 359
Dh 1157 CCGGAG 1216
Qy 360 LeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAla 379
Dh 1217 TTGGAAG 1273
Qy 380 ThrGluValPheAspGluLysMetGluAlaHisGluGluValValAlaGluValAlaHis 399
Dh 1274 ACCGAAGGTGTTATATAG 1333
Qy 400 SerThrValGluLysThrGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 419
Dh 1334 AGCAGCGTGAAGAGATGACAGAAAGGAGCA---GAGAGAGAGAGAGAGAGAGAGAGAG 1390
Qy 420 ValValGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 439
Dh 1391 GTGGTGAAGAGATCGGAGAGATCTTGGCCCTCGAAGAACTGGCTGAGACCCGAGAGGTC 1450
Qy 440 ProGlnGluAlaGluProAlaGluGluLeuMetLysSerAlaArgGluMetCysValSerGly 459
Dh 1451 CCCCAGAGAGCTGAGCCCTGTGAGAGAGCTGATGAAGAGCAAAAGAAAGTATGCTCTGGG 1510
Qy 460 GlyAspHisThrGluGluLeuThrAspLeuSerProGluGluLysThrLeuProLysHisPro 479
Dh 1511 GGTGACCACTACAGCTGACAGATCTAAAGTCTTAAGAGAGAGAGAGAGAGAGAGAGAG 1570
Qy 480 GtngluIleValSerGluValAlaGluMetLeuSerSerGlnGluAlaGlyValGlnGly 499
Dh 1571 GAAGGCATTTGTCAGTGAAGTGAAGTGTGCTGCTTCCAGAGAGAGATCAAGATCAGGGA 1630
Qy 500 SerProLeuLysLysLeuPhePheSerSerSerGlyLeuLysLysSerGlyLysGln 519
Dh 1631 AGTCCCTGAAGAGAGCTTCTTCAAGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1690
Qy 520 LysGlyLysArg---GtngluGlyGlyAspGluGluGluGluGluGluGluGluGluGluGlu 538
Dh 1691 AAGGAG 1750
Qy 539 ThrGluSerProGluSerAlaAspGluGluGluGluGluGluGluGluGluGluGluGlu 558
Dh 1751 ACCAGAGTCCCGAG 1810
Qy 559 GtngluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGluAspGly 578
Dh 1811 GAGCCCGAGAGAGATGCGCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1870
Qy 579 AlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 598
Dh 1871 GCTGAG 1930
Qy 599 SerPheLysLysMetValThrProLysLysArgValAlaArgProSerGluSerAspLys 618
Dh 1931 TCCTTCAAAAGATGTGTACACACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1990
Qy 619 GtngluGluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSer 638
Dh 1991 GAAGAGAGAGCTGATGAAGTCAAGAGTCCACCTTGTCTCCACAGAGAGAGAGAGAG 2050
Qy 639 GtngluMetGluAspGluValLysThrValGlyGluGluGluGluGluGluGluGluGlu 658
Dh 2051 GGAATGACAGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2110
Qy 659 ArgValAspThrSerValSerThrGluAlaLeuIleCysValGlySerSerLysLysArg 678
Dh 2111 AGGTGTGATCTTCACTGCTCTTGGAGGCGTTGATTTGTGCGATCGTCCAAAGAGAG 2170
Qy 679 AlaArgLysAlaSerSerSerSerAspGluGluGlyProArgThrLeuGlyLysAspSer 698
Dh 2171 GCGAGAGAGAGCATCTCTTCAAGATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2230
Qy 699 HisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer 718
Dh 2231 CACAGAGCGGAG 2284
Qy 719 ThrGluGluGluAspGluAlaGluGluSerSerSerProGluProAlaGlySerProSer 738
Dh 2285 ACCCAG 2344
Qy 739 GtngluGluGluValSerThrThrProLysPheLysArgLeuValThrProArgLysLys 758
Dh 2345 GAAGGAG 2404
Qy 759 SerLysSerLysLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 778
Dh 2405 TCCAGTCAAAATCGAG 2464
Qy 779 IleGluProSerArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 798
Dh 2465 ATCGAAGCAAGTGAAGAGAGATCTTGGCTTCCATTAAAGAAATTTATCTCTGAGAGCGG 2524

QY 799 LyeIysArgAlaAspGlyLysGlnGluAlaThrValGluAspSerGlyProValGlu 818
DB 2525 AAAAAAGGCGAGATGGAGACAGAAACAGCGCCCGCTTTGAAGACTCGGGGCCAGGAG 2584
QY 819 IleAsnGluAspAspProAsnValProAlaValValProLeuSerGluTyrAsnAlaVal 838
DB 2585 ATCAATGAGGACGACCCCGACGCTCCAGCTGTGTGCTCTGCTGAGTAGATGCGGTA 2644
QY 839 GluArgGluLysMetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeu 855
DB 2645 GAGAGAGAAGAGCTGGAAAGCGCAGCGAGCTCAGGAGAACGTGGAGCTGCCAGCTGAAG 2704
QY 856 GlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThrValSerValAla 875
DB 2705 GGGGCTGTGTATGTCTGAGGAGCTTAGTAAGACTCTGGTTACACCGTAGTGTCCGG 2764
QY 876 ValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIleSer 895
DB 2765 GTCAATTGATGGACACGCGCAGTCACCAAGTCCGGAAGCGGTCCCTTCTGTGGATATCT 2824
QY 896 AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGlu 915
DB 2825 GCTTCCATGACAGAACTCTTGTAGCACGACGAGGGAGTGGCCACACCGCTGTGTGGAGAG 2884
QY 916 ValThrGluLysAspIleIleAlaGluGluThrProValLeuThrGlnThrLeuProGlu 935
DB 2885 GTCACTGAAAAGACATCATCTGCAAGAACACTCTCCATCTCGCCCACTTTACCAAGG 2944
QY 936 GlyLysAspAlaHisAspAspMetValThrSerGluValAlaAspPheThrSerGluAlaVal 955
DB 2945 GGCAGAATGCCATGACGACATAGTCCACAGTGAAGTGAATTTTACTCTAGAAGCAGTG 3004
QY 956 ThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAla 975
DB 3005 ACAGCCGACAGAAACACAGAGGCGCTCCGCGCTGAAGAACTTACCGAAGCATCAGGGGCA 3064
QY 976 GluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThr 995
DB 3065 GAAGAGACCCACAGACATGTGTCTGCAAGTTTCCAGAGTGTCCGACTCCCGCGACACACA 3124
QY 996 GluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGlu 1015
DB 3125 GAGGAAGCCACCCAGTTTCAAGAGTGAAGGTGAGAGTGGCATGTAGATACGGAAGAACAGGAG 3184
QY 1016 ArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnVal 1035
DB 3185 CGCCAGACGAGCGCGTCTCCAGCGCTTGCACAGAAAGTGAAGAGGACTCCAGGTG 3244
QY 1036 ProAlaThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluVal 1055
DB 3245 CCTGCAACCCAGACTCTGCAGAGACGAGGCGCGAAAGCACTGGAGAGGTGGAGAGGTA 3304
QY 1056 GluGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyPro 1075
DB 3305 GAGAGAGACTCCGAGGTGTGGCTTACCAGAAAGAGAAGGATGTGTGCGGAAGGACCC 3364
QY 1076 ValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrPro 1095
DB 3365 GTGCAGGAAGCTGAAGTGAAGTACAGAGGATGTAGACGCTCCGAGACTGTACAGGCTACCCCA 3424
QY 1096 GluSerLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnVal 1115
DB 3425 GAGAGCCTTGAAGTTCCTGAAGTACAGAGGATGTAGACGCTGCCACCATGCGCAGGTT 3484
QY 1116 IleLysLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThr 1135
DB 3485 ATCAAGCACCAGCAGCTGTATGGAAACAGGCTGTGGCCCTGAGTCATCTGAACCTTGACA 3544
QY 1136 AspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln 1155
DB 3545 GACAGTGAGACAAATGAAGTACTCCCTCGCAGATTACAGACTCCAAACGGGACACAG 3604
QY 1156 GlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaValArgGlnSerGln 1175

DB 3605 CAAGACGAGACCGTTTACAGCCAGGACAGTAATGCCATTGCGCCCGCTCAAGCAGTCACAG 3664
QY 1176 ValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsn 1195
DB 3665 GTCACTGAAGAGGAGGACGCTGCTGTCTCAGCGAGGGGCTTCAACA---CCATCTAGT 3721
QY 1196 ValProAlaGlnGluGluHisGlyGluGluProGlyArgAspValLeuGluProThrGln 1215
DB 3722 TTTCCAGCCCGAGAAACACACAGGAAAAACAGGAAGGATGTTCTAGAACCCACA--- 3778
QY 1216 GlnGluLeuThrAlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGly 1235
DB 3779 CAAGCCCTGCTGCTCGGGGAGTGCCTATTCTGGCAAGGCTGAGGTGGTCAAGAGGT 3838
QY 1236 GluValAspTrpLeuAspGlyGluLysValLysGlu-----GluGlnGlu 1250
DB 3839 GAGCTGCCCATTTGATGGAGAAAAGTCAAGACGGACAGTGTGTTAAAGAACTGGAG 3898
QY 1251 ValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGlu 1270
DB 3899 GTGCTGTGTGCACACTGGACCCCAACAGTCAAAAGACTGCTGACTTGACACGTGACAGTAA 3958
QY 1271 ValMetClyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSerLeuSerLeu 1290
DB 3959 GTAATGGAAGTGGCCAGATGTCTAGGAAACTGAGAGTAATGAAGAACAGAGTATTAGCCG 4018
QY 1291 GluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGluGln 1310
DB 4019 GAGAAAGAGAGATGGGAACCCGACGTTGAAAGAGGAGAAACAGACAGACAGAGCAA 4078
QY 1311 ValSerGluGluGlyGluGlnGluThrAlaAlaProGluHisGlyThrTyrGlyLys 1330
DB 4079 GCCAGTGAAGAAACATCAGCAGGAAACAGCTGCTCTCTGAGCATGAAGGAACCCACCTAAG 4138
QY 1331 ProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGly 1350
DB 4139 CCAGTCTCTCACAGCTGACATGCTCTCCTCAGAGAGGGAAGGACACTGGGACGCTTGA 4198
QY 1351 GlySerProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSer 1370
DB 4199 GGAAGCCCTTCTCTCCAGACCCAGCAAGCAAGCAGATTGATAGAGTTCAAGTTCAAAGC 4258
QY 1371 LeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValIleGluThrValVal 1390
DB 4259 TCAGACACACCACTCACTCAAAACAACCGAAGCTGTGAAAAAGTCCGAAGAACTGTGGCA 4318
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
DB 4319 ACTTCAGAGATGGATGAAAGTTTGGAGTGTGCAGGTGCGCAATCATTTACCAGCTGAGAAG 4378
QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
DB 4379 CTCCTCGAAACCGGTGGCTACGGGACTTTCAGCATGGAGAGGACACCTGTGCCCGAGGG 4438
QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
DB 4439 CCTGAGTCTCAGCGACAGTCCATCCCAATAATAGTAACCTCTCTCTCTGAAAGCATCCTA 4498
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
DB 4499 CATTCAGCCTTCAAGAGAGAAGTGAAGCATCCCAAGAAACAGAGATCAGATGAAGATAAC 4558
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
DB 4559 AAGCCAGATGCTGGTCTCTGATGCTCCGCAAGGAGAGTGCAGAGAGAGAAAAATCCTC 4618
QY 1491 LysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIleValLeuAsnVal 1510
DB 4619 AGGGCTGAACCTCAGATCTTGGAACTTGAGATGAAGCAATGAAGATTGTCCAGAGTGT 4678
QY 1511 IleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHisAlaTyr 1530

Db 4679 ATCCGACAGCCGTCGACCGATTGTGACCTACAGAAACGCCCCGGAACCCACGCTTCT 4738
Qy 1531 AspSerGlnThrGlnValProAla-CysArgLeuAspSerArgGluProAsnArgCysTr 1550
Db 4739 GATTATACAGAAATCAGATCTCTGTGATGACGCTGACGACGAGGACACAAACAGATGCTG 4798
Qy 1550 PhrIlyMetIys-----AspAlaIlyMetIysShIleProValProGlnIrr 1565
Db 4799 GACAAAGATGAAAGACGACCTTCAAGTCTCCCGCAAGATGAAACACTCAGTCCCGTAGCC 4858
Qy 1565 oArgGluAspLeuGlnValIleuThrValIleuGlnIuaIaTProAlaGlnProArgIlyCysIle 1585
Db 4859 CAGGAAAGACTTTCGCGTTCTGTATAGTTCGTGAAGGATAGCAAGGCTTCAGAAATGATC 4918
Qy 1585 uProArgLeuGlnIleuIlyAspAlaProValSerIlyS 1596
Db 4919 ACCAGCGTTCGAGTTGAAAGTGCCAGCTGTCAAG 4952
RESULT 5
AF326228 6197 bp mRNA linear ROD 16-JUL-2001
LOCUS AF326228
DEFINITION Mus musculus PKC binding protein SSeCKs mRNA, complete cds.
ACCESSION AF326228
VERSION AF326228.1 GI:13236403
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 6197)
Camus, R., Mesbah, K., Rallu, M., Babinet, C. and Barra, J.
Gene trap insertion reveals two open reading frames in the mouse
SSeCKs gene: the form predominantly detected in the nervous system
is suppressed by the insertion while the other, specific of the
testis, remains expressed
Mech. Dev. 105 (1-2), 79-91 (2001)
2132283
11429284
2 (bases 1 to 6197)
Camus, A., Mesbah, K., Babinet, C. and Barra, J.
Direct Submission
Submitted (06-DEC-2000) Immunologie, Institut Pasteur, 25 rue du
Docteur Roux, Paris 75015, France
Location/Qualifiers
1..6197
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="10"
140..5194
/codon_start=1
/product="PKC binding protein SSeCKs"
/protein_id="AAK16150.1"
/translation="MGAGSSTFRSPEOPASDTPSELESGHPSAASGAAGPAD
ADPATLPOKQQLSANGVAQOEYHVOEESQDOEEVYTEDQGESSEVKKDR
AEKMAASSTVVEDITKDEEETPEITIEQIPASEVVEEAQIAESQANDVKKAFK
VGFTKVKDKNEKSDTVQLTVKKDVEGASVAGVAGGKEMNEKPTPLKLEPTPS
KOSTKEOEGTLKQAOASTEIPLOAESGGQTEBEAAKADGEMNEKPTPLKLEPTPS
NETTSFPKFTLHGWAAMRKSTSPFKPKEDLETSEBKAEQAKVDEEKEETPAP
AEORPAGCTDQARLSADYEKYLEPLEDQGVLELSEKCALATEVDEPREKHAOEY
VAEVSTYERKTKGQGAEEVGDVYVESGESLPEKLAETQEVPOEAEPEBELMKT
KEVCSGDHQLTLDLSEPKMLPHPEBIVSEVENLSSQERIKVQSPFLKLFEISSG
LKLKSKQKQKGRGGGGGDEBEYQHITGESPEASDEBQGSASPEPEELICLE
KSPAPQOAEAEBAEGATSDGEKKEGITPMASFKMVTPEKKVRPSESKKEELDKY
KSATLSTSTESGMODEYRAVGEORSEPKRYDTSWEALICVSGSKRARKAS
SSDDGSPRTTIGDGHRAEAKDKKADLPASTOBODAHSSSSPEPAGSPSEGEY
STWESFKRLVTRFKSKSKLEBAEDSGAEQLASIEPBERESWISIKKFTIGRKKR
ADGQOQAEVDSGPEINEDPDVPAVPLSEYDAVEREKLAEQRAQENVLPOLK
AVYBEELSKTLVHTVSVAVIDGTRAVTSAEBSRPSMTSASMTPELHAEGVATPVG
EYTEKDTAEATPALAQLTGGKADHDIVTSEVPTSAVVAETTELALREELTEA
SGAEFTTMVVASLOSPTTEATPVQEVGGMLDTEBEROTQAVLOAVADKVK
EESQVATOTTLQAGKALEKVEYBESEVALTEKRVNVEGPGQAEETHLQGS
ETVQATPESLEVEYTEDVDRAITTCVVIKHQQLMEQAVAPSSSETLITDSENGSTPLA

DSPTNGTODDETVDSDSNALIAVKOSVTEBEAAAOTESPSSPPAOEHRK
PKRDVLEPTOTALAAGAPILAKAEVQEBGACGPDEKVDGCVLELVPYHTGPS
QKADVLRTDSVNEVARCQETESNEEDSISPEREKWTDVKEKETETKTEQASBEHBQ
ETAPABEGHTRPKPVLTADMPHSEBGRALGSLGSPBLPOQDADCLEVOVQSDTPV
TOTTEAVKKEEVAVTASEMDESLGCAQSLPAEKISETGYTLQIGEDTVPOGPS
OASIEPIVTPAPESILHSDLORESVASKORSEDEBKPAGPDAKESAAEKLIR
ADPEILEESKSNKIVSVIOTVADQARTETAPETHASLONOVPMQADSGAOM
LDKDESLQVSRDQGTLSAVAOBAGVANSDEKSKRASEMTTLTAVESAVKSEVEKL
PLOCKDEKHPHPOHOSLAKAEDASGNLTIXESPDTNPKLTFEBDGLKEKMKAO
TEEDDLOEPPKDLTES"
BASE COUNT 1840 a 1447 c 1779 g 1131 t
ORIGIN
Alignment Scores:
Pred. No.: 7.31e-230 Length: 6197
Score: 6663.00 Matches: 1357
Percent Similarity: 88.90% Conservative: 76
Best Local Similarity: 84.18% Mismatches: 156
Query Match: 82.53% Indels: 23
DB: 10 Gaps: 11
US-09-902-432-4 (1-1596) x AF326228 (1-6197)
Qy 1 MetGlyAlaGlySerSerThrGlnGlnArgSerProGlnGlnProAlaGlySerAspThr 20
Db 140 ATGGGTGACGACGATTCACCGAGCAGCGAGCCCGGACGACCGCGGAGACGACACG 199
Qy 21 ProSerGlnLeuValLeuSerGlyYhiGlyProAlaIaGlnIaIaSerGlyValAlaGly 40
Db 200 CCGACGACGCTGAGACTCAGTGGCCATGGCCCGGACGGAAGCGTCCGGAGACGACTGA 259
Qy 41 AspProAlaAspAlaAspProAlaThrIlySleuProGlnIlyAsnGlyIghIleuSerSer 60
Db 260 GATCCCGCTGACGCGGACCCGCCACCAAGCTCCACAGAAAGATGTCAGCTGTCTCC 319
Qy 61 ValAsnGlyValAlaGlnGlnGlyIlyAspValIhSvalGlnGlnGlnIlyGln 80
Db 320 GTCAATGGTGTAGCTGAACCAAGAGATGTCCAGTCCAGGAAGGACGAGATGGGCA 379
Qy 81 GlnGlnGlnValIaIaSpGluAspValGlyGlnArgGlnSerGluAspValArgGlnIlyS 100
Db 380 GAGGAAGAAGTCACTGTTGAAGATGTTGACAGAGAGTCAAGAAAGATGTGAAGAAAA 439
Qy 101 AspArgValIaGlnGlnMetAlaIaAsnSerThrIaValIaGlnAspIleThrIlyAspGly 120
Db 440 GACCCAGCTTAAAGAAATGACGCCAGCTTCCACAGTTGTGAAGATATCAAAAGACGAG 499
Qy 121 GlnGlnGlnIlyThrSerGlnIleIleGlnGlnIleProAlaSerGlnIlyAsnValGlnGln 140
Db 500 CAGGAGGAAACACCGGAATTAATCAACAGATCCCTGCTTCAGAGGACAAATGTGAAGAA 559
Qy 141 MetValGlnProAlaGlnSerGlnIaIaAsnAspValGlyPheIlySlyValPheIlySph 160
Db 560 ATGGCGACGCGTGTGAGTCCCAAGCTAATGACGTCCGCTTCAAGAAAGTATTAAATTT 619
Qy 161 ValIglyPheIlySphThrValIlySlyAspIlyAsnGlnIlySereAspThrValGlnIleu 180
Db 620 GTTGTGTTTAAATTACCGGTGAAGAAAGATTAAGAAAGATCGATCAAGCTCA 679
Qy 181 LeuThrValIlySlyAspGlnGlnGlyIlyGlnGlnIaGlnIaIaSerValGlyValAspIhS 200
Db 680 CTCACGTGCAAGAAAGATGAAAGGCGCAAGGCGCAAGGCTCCGTTCGAGCAGAGAACAC 739
Qy 201 GlnGlnProSerValGlnIlyThrIaValGlyIlyGlnIlyGlnIlyGlnIlyGlnIly 220
Db 740 CAAAGCCCGGAGTGAAGACC---GTCCGCGAATCAGCATCCAAAGAAAGTACGTGAAG 796
Qy 221 GlnSerThrGlnIlyGlnGlnIlyThrIleuIlyGlnGlnIlyGlnIlyGlnIlyPro 240
Db 797 CAATCCACAGAAACCAAGAGGACCCCTGAAGCAAGACAGACAGACAGAAATTTCC 856
Qy 241 LeuGlnIaIaGlnSerAspGlnIaIaGlnGlnGlnIlyIlyAspGlnGlnIlyGlnIlyS 260

Db 857 CTTCAAGCCGAATCTGGTCAAGGACCGAGGAAGACGCCAAAGATGGAGAAGAAAC 916
Qy 261 GlnGluLysGluProThrLysSerProGluSerProSerSerProValAanSerGluThr 280
Db 917 CGAGAGAAAGAACCTACCAAGCCCTAGAACTCTCCGACAGCCCTGTTCAGCAATGAGACA 976
Qy 281 ThrSerSerPheLysLysPheThrHisGlyTrpAlaGlyTrpArgLysLysThrSer 300
Db 977 ACATCTTCTTCAAGAAATTTCTTCACTCAGGGTGGCCGCTGGCCGCAAGAACCCAGC 1036
Qy 301 PheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAla 320
Db 1037 TTCAAGAAACCAAGGAAGATGATCTGAAACTTCCGAGAGAGAAAGAGCAGAGGCT 1096
Qy 321 GluLysValAspGluGluLysGluLysThrGluProAla---SerGluGluGlnGlu 339
Db 1097 GAAAAAGTAGACGAGGAAGAGGGGAAAGACAGAGCCAGCCAGGAGGAGGAG 1156
Qy 340 ProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuPro 359
Db 1157 CTTGAGAGGACAGACACAGCCAGGCTGTTCAGCCGACTATGAGAAAGTGGAGTTGCCT 1216
Qy 360 LeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAla 379
Db 1217 TTGGAAGACCAAGTCTGGTACCTGGAGGCATTGTCTG---GAGAAAGTGTCTCTTTGGCA 1273
Qy 380 ThrGluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisVal 399
Db 1274 ACGGAAGTGTGATGAGAAGACGGAAGCCACCAAGAAAGTGTGTTGAGAGGTTCCACGTG 1333
Qy 400 SerThrValGluLysThrGluGluGlnGlnGlyGlyGlyGluAlaGluGlyGlyVal 419
Db 1334 AGCACCGTGAGAAAGATGACGAAGAGGGCA---GGAGGACAGAGGTGGAAAGGGGATGTG 1390
Qy 420 ValValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluVal 439
Db 1391 GTGGTGAAGATCGGAGAAATCTTCTCCCTGAGAAATCTGCTGAGACCCAGGAGTCT 1450
Qy 440 ProGlnGluAlaGluProAlaGluLeuMetLysSerArgGluMetCysValSerGly 459
Db 1451 CCCAGGAAGCTGAGCTGTGGAGGAGCTGATGAAGACCAAGAAAGTATCGCTCTCTGGG 1510
Qy 460 GlyAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisPro 479
Db 1511 GGTGACCATACTACGTGACAGATCTAAGTCTTGAAGAGAGATGCTTACCCAAACACCCC 1570
Qy 480 GluGlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGly 499
Db 1571 GAAGGCATTTGTCAGTGAGTGGAGATGCTGCTCTCAGAGAGAAATCAAGGTACAGGGA 1630
Qy 500 SerProLeuLysLysLeuPheSerSerGlyLeuLysLysLeuSerGlyLysLysGln 519
Db 1631 AGTCCCTGAAGAAGCTCTTTCAGCAGTTTCGGGCTTAAAGAAGCTCTCCGGGAAGAGCAG 1690
Qy 520 LysGlyLysArg---GlyGlyGlyAspGluProGlyGluTyrGlnHisIleHis 538
Db 1691 AAGGGGAAGAGAGAGGAGCGGGGAGATCAAGAGCCAGGAGAAATACCAACACATTCAA 1750
Qy 539 ThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGlu 558
Db 1751 ACCGAGTCCCGAGAGTGTGACGACAGAGGAGGAGAGTCTGCTCTTCCCTCGAA 1810
Qy 559 GluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGlu 578
Db 1811 GAGCCCGAGGAGATCGCGTGTCTGGAGAGGGGCCCTCGGAAGCACCCCAAGAGCGGAA 1870
Qy 579 AlaGluGluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProThrAla 598
Db 1871 GCTGAGGAGAGGAGGCTTCCGACGAGAGAGAAAGGAGGAGGATCACCCCTGGCA 1930
Qy 599 SerPheLysLysMetValThrProLysLysArgValArgArgProSerGluSerAspLys 618
Db 1931 TCCTTCAAAAGATGGTGAACACCAAGAACCGGTCCGAAGACCTTCTTGAGAGCGCAAG 1990

Qy 619 GluGluGluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSer 638
Db 1991 GAAGAAGAGCTGGATTAAGGTCAAGAGTCCACCTTCTCTCCACGAGAGCACGGCTCT 2050
Qy 639 GluMetGluAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArg 658
Db 2051 GGNATCGAGATGAGTTCAGCGGTTGGCAGGAGCAAGAGTCAGAGAGCCAAAGCGC 2110
Qy 659 ArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArg 678
Db 2111 AGSGTGGATACCTTCAGTGTCTTGGAGGGCTTGAATTGTGTCGATCGTCCAAGAGAGA 2170
Qy 679 AlaArgLysAlaSerSerSerSerAspGluGlyGlyProArgThrLeuGlyGlyAspSer 698
Db 2171 GCAGAGAAAGCATCTCTTCAGATGATGAAGGAGGCCAAGAACACTGGGAGGGGATGGC 2230
Qy 699 HisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer 718
Db 2231 CACAGAGCGGAGGAGGTAGCAAAAGCAAAAGAGCA-----GATGCTCTTCTTCGCCAGC 2284
Qy 719 ThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSer 738
Db 2285 ACCCAGAACAGACCAAGCGCACGGAAGTCTCTCACCCGAGCCAGCTGGAAGCCCTTCT 2344
Qy 739 GluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLys 758
Db 2345 GAAGGCGAGGGGCTCTCCACCTGGAGTCAATTAAGAGATTAGTACTCCACGAAAAAAA 2404
Qy 759 SerLysSerLysLeuGluLysAlaGluAspSerSerValGluGlnLeuSerThrGlu 778
Db 2405 TCCAAGTCAAACTGGAAGAGAGAGCCGAGAGTCCCGGTGCAGAGCAGTTGGCCTCCGAG 2464
Qy 779 IleGluProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArg 798
Db 2465 ATCGAACCAAGTAGAGAGGAATCTTGGGTTTCCATTAAAGAAATTTATTCTTGACCGCGG 2524
Qy 799 LysLysArgAlaAspGlyLysGlnGluAlaThrValGluAspSerGlyProValGlu 818
Db 2525 AAGAAAGCGCAGATGGGAAGCAAGAACAGGCGCGCTTGAAGACTCGGGGCCAGGAGAG 2584
Qy 819 IleAsnGluAspAspProAsnValProAlaValProLeuSerGluTyrAsnAlaVal 838
Db 2585 ATCAATGAGAGAGACCCCGACGCTCCAGCTGTGTGCTCTGTGAGTACGACGCGTA 2644
Qy 839 GluArgGluLysMetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeu 855
Db 2645 GAGAGAGAGAGCTGGAAGCGCAGCGAGCTCAGGAGACCTGGAGCTGCCACGCTGAAG 2704
Qy 856 GlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThrValSerValAla 875
Db 2705 GGGGCTGTGTATGTCTCAGGAGCTTAGTAAGACTCTCGTTTCCACTGTGAGTGTGCG 2764
Qy 876 ValIleAspGlyThrArgAlaValThrSerValGluArgSerProSerTrpIleSer 895
Db 2765 GTCAATTGATGGACGAGGCGAGTCAACGTCGCAAGAGCGGTCCCTTCGTGGATATCT 2824
Qy 896 AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGlu 915
Db 2825 GCTTCATGACAGAACTCTTGAAGACCGAGAGGAGTGGCCACCGCTGTGTGAGAGAG 2884
Qy 916 ValThrGluLysAspIleAlaGluGluThrProValLeuThrGlnThrLeuProGlu 935
Db 2885 GTCACTGAAAAGACATCACTCAGAGAGCAACTCTCTGCACTCGCCAGACTTTACCAAGG 2944
Qy 936 GlyLysAspAlaHisAspAspMetValThrSerGluValAspPheThrSerGluAlaVal 955
Db 2945 GGCAGAGATGCCCATGACGACATAGTCAACAGTGAAGTGGATTTTACCTCAGAAAGCAGTG 3004
Qy 956 ThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAla 975
Db 3005 ACAGCGCAGAAACACACAGAGGCGCTCCGCGCTGAAGAACTTACCGAAGCATCAGGGGCA 3064

QY 976 GUGUUTThrThraspMetValSerAlaValSerGlnLeuThraspSerProAspThrThr 995
 Db 3065 GAAAGAGACACAGACATGCTGCTGCACTTCCAGACTCCCGAGCACACACAA 3124
 QY 996 GUUGUUAAThThProValGlnGlnValGlnSerGlyValLeuAspThrGlnGlnGln 1015
 Db 3125 GAGGAAGCCACCCCACTTACAGAGAGTGAAGGGGTGCATCTGATACGGAAGAACAGAG 3184
 QY 1016 ArgGlnThrGlnAlaLeuGlnAlaValAlaAspValValGlnGlnSerGlnVal 1035
 Db 3185 CCCAGACCCAGGCGCTCTCCCAAGCGTTGCGAGCAAAAGTGAAGAGCACTCCAGGTG 3244
 QY 1036 ProAlaThrGlnThrValGlnArgThrGlySerValAlaGlnGlnValGlnGlnVal 1055
 Db 3245 CCGCAACCCAGACTCTGCGAGAGAGAGGCGCCGAAAGCACTGAGAGAGGAGGAGGTA 3304
 QY 1056 GUGUUAAspSerGlnValLeuAlaSerGlnValGlnValAspValMetProGlyGlyPro 1075
 Db 3305 GAGGAGAGACTCCGAGGTCTGCTTACCGAAGAAAGAGAGATGTTGTCGCGAGAGACC 3364
 QY 1076 ValGlnGlnValGlyValGlnGlnValGlnGlnGlnGlnGlnGlnGlnGlnGln 1095
 Db 3365 GTCGAGAGAGCTGAAACTGAGCATCTTGCAAGGCGCTCCGAGACTGACAGGCTACCCA 3424
 QY 1096 GlnSerLeuGlnValProGlnValThraAlaAspValAlaAspValAlaThraCysGlnVal 1115
 Db 3425 GAGAGCTTGAAGTCTTGAAGTCAAGAGAGATGAGACCGTCCACCAATGCCAGGTT 3484
 QY 1116 IleValLeuGlnGlnLeuMetGlnGlnAlaValAlaProGlnSerSerGlnThraLeuThr 1135
 Db 3485 ATCAAGCACACAGAGCTGATGAAACAGCTGTGCCCCGTGAGTCACTGAAACCTTGACA 3544
 QY 1136 AspSerGlnThraAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln 1155
 Db 3545 GACAGTGAAGCAAAATGCAAGTACTCCCTCGAGATTGACACACTCCAAACGGGAGACAC 3604
 QY 1156 GlnAspGlnThraIleAspSerGlnAspSerValAlaThraAlaValAlaArgGlnSerGln 1175
 Db 3605 CAGAGAGAGCCGTTGACAGCCAGGACAGATGATGCAATGCCCCGCTGAGCAGTCAAG 3664
 QY 1176 ValThrGlnGlnGlnAlaAlaThraAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 1195
 Db 3665 GTACTGAAAGAGAGGAGGAGCTGCTGCTCAGACGAGGGGCTTCAACA---CATCTGAT 3721
 QY 1196 ValProAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1215
 Db 3722 TTTCCAGCCCGAGAGACACACAGAGAAACACAGAAAGGATGTTCTAGAACCCACA--- 3778
 QY 1216 GlnGlnLeuThraAlaAlaValAlaProValLeuAlaGlySerThrGlnValGlnGlnGln 1235
 Db 3779 CAGAGCCCTGCTGCTGCGGAGAGTCTTATCTGCAAAAGGCTGAGGTGAGGAGGCT 3838
 QY 1236 GlnValAspTrpLeuAspGlnGlnValValGlnGlnGlnGlnGlnGlnGlnGlnGln 1250
 Db 3839 GAGGCTGCGCCAGTTTATGAGAAAGTCAAGACGAGACAGTGTGTAAGAACTGAGAG 3898
 QY 1251 ValPheValIleSerGlyProAsnSerGlnValAlaAlaAspValThraAspSerGln 1270
 Db 3899 GTGCTGTGACACTGACCAACAGTCAAAAGACGTGCTGACACTTGAACAGTGAAGAA 3958
 QY 1271 ValMetGlyValAlaGlyCysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1290
 Db 3959 GTATATGAGATGCGCAGATCTCAGAAACTGAGATGATGAAACAGATATTAACCCCG 4018
 QY 1291 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1310
 Db 4019 GAGAAAGAGAGAGTGGAGAACCGACTTGAAGAAAGAGAAACAGAGACCAAGAGAGCAA 4078
 QY 1311 ValSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1330
 Db 4079 GCCAGTGAAGAAATGATGAGAGAAACAGTGTCTGAGCATGAAAGAAACCAACCTTAAG 4138
 QY 1331 ProValLeuThraLeuAspMetProSerSerGlnArgGlyValAlaLeuGlnSerLeuGly 1350

Db 4139 CCAGTCTGACAGCTGACATCTCCTCACAGAGAGGAGAAAGCACTGGAGACCTTGAA 4198
 QY 1351 GlySerProSerLeuProAspGlnAspValGlyCysIleGlnValGlnValGlnSer 1370
 Db 4199 GGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGATTGATGAGAGTTCAAGTCAAGC 4258
 QY 1371 LeuAspThrThrValThrGlnThraAlaGlnValAlaGlnValValIleGlnThraVal 1390
 Db 4259 TCAGACACACAGATCACTCAACCAAGCAGAGCTGGAAGAAAGTGAAGAACTGTGCA 4318
 QY 1391 IleSerGlnThrGlnGlnSerProGlnCysValGlyAlaIleIleLeuLeuProAlaGly 1410
 Db 4319 ACTTCAGAGATGATGAGAAAGTTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4378
 QY 1411 SerSerAlaThrGlnGlyIleIleIleIleIleIleIleIleIleIleIleIleIle 1430
 Db 4379 CTCTCCGAAACCGGTGCTGAGGAGCTTCCAGATGAGAGAGACACCGTCCCGAGGGG 4438
 QY 1431 ProGlnSerGlnAlaGlnSerIleProIleIleValThraProAlaProGlnSerThrLeu 1450
 Db 4439 CCGAGTCTCAGGAGAGATCCCAATATGATTAATCTGCTCTGGAAGACATCCTTA 4498
 QY 1451 HisProAspLeuGlnGlnGlnIleSerAlaSerGlnArgGlnArgSerGlnGlnGln 1470
 Db 4499 CATTTGACCTTCAAG 4558
 QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGlnSerThrAlaIleGlnValLeu 1490
 Db 4559 AAGCCAGATGCTGTGCTGATGCTGCGGCAAGAGAGATGACAGCAAGAGAAATCTTC 4618
 QY 1491 LysAlaGlnProGlnIleLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1510
 Db 4619 AGGCTGAACCTGAGATCTGGAACCTGAGATGAGATGAGATGAGATGAGATGAGATG 4678
 QY 1511 IleGlnThraAlaValAspGlnPheAlaArgThrGlnThraAlaProGlnThraIleAla 1530
 Db 4679 ATCCAGACAGCCGTGACACAGTTCAGATGACAGAAACAGCCCGGAAACCAAGCTTCT 4738
 QY 1531 AspSerGlnThraGlnValProAla-CysArgLeuAspSerArgGlnProAsnArgCysTr 1550
 Db 4739 GATTATACAGAAATCAGATCTCTGTGATGACAGCTGACAGCGAGGAGCAACAGATGCTG 4798
 QY 1550 PthIlyMetLys-----AspAlaLysMetLysHisProValProGlnPr 1565
 Db 4799 GACAAAGATGAAAGGACCTTCAAGTCTCCCCCAAGATGAAACCTCAGTCCCTTACC 4858
 QY 1565 CArgGlnAspLeuGlnValIleThraValIleGlnGlnAlaTrpAlaGlnProArgLysCysLe 1585
 Db 4859 CAGGAAGATCTGCGGCTTCTGATGATGTTCTGAAAGCATGAGCAAGGCTTCAAGATGATC 4918
 QY 1585 uProArgLeuGlnLeuLysAlaProValSerLys 1596
 Db 4919 ACCACGCTTGACAGTTGAAAGTCCAGTGTCAAG 4952

RESULT 6
 AF326230 5902 bp mRNA linear ROD 16-JUN-2001
 LOCUS Mus musculus testis-specific SSecks mRNA, complete cds.
 DEFINITION AF326230
 ACCESSION AF326230
 VERSION AF326230.1 GI:13236407
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 5902)
 REFERENCE
 1 Camus,A., Mezbah,K., Rallu,M., Babinet,C. and Barra,J.
 Gene trap insertion reveals two open reading frames in the mouse
 SSecks gene: the form predominantly detected in the nervous system
 is suppressed by the insertion while the other, specific of the
 testis, remains expressed
 Mech. Dev. 105 (1-2), 79-91 (2001)

[illegible]

QY	528	AspGluGluProGlyGlyValTyrGlnHisIleHisThrGluSerProGluSerAlaAspGlu	547
Db	1423	GATGAAGGCCAGGAAATACCAACATTCAACCGAATCTCCCAAGAGTGTGACAG	1482
QY	548	GlnYrsGlyGluSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGlu	567
Db	1483	CAGAAAGGAGAGACTCTGCTCTTCCCTCGAAGAGCCCGAGAGATCGCGTCTGGAG	1542
QY	568	LysGlyProLeuGlnAlaProGlnAspGlyGlnAlaGluGluGlyThrThrSerAspGly	587
Db	1543	AAGGGCCATCGAAGACACCCCGAAGACCGAAGCTGAGAAAGACCACTTCCGACGGA	1602
QY	588	GlnYrsYsAspGluGlyIleThrProTrpAlaSerPheYsYsMetValThrProlys	607
Db	1603	GAGAAAGAAAGGGAAGGATACCCCCCTGGGATCTCTTAAAAGATGGTAGCACCCAAAG	1662
QY	608	LysArqValArgArgProSerGluSerAspYsGluGluGluMetGlnYsValYsSer	627
Db	1663	AAACGGGTCGCAACACTTCTGAGCGCAAGAAAGAAAGAGCTGATAGGTCAAGACT	1722
QY	628	AlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValYsThrVal	647
Db	1723	GCCACTCTTGTCTCTCAGAGAGAGCAGCGCTCTGGAATCAGAGATGAGGTCAAGCGTTT	1782
QY	648	GlyGluGluGlnYsProGluGluProYsArgYsValAspThrSerValSerTryGlu	667
Db	1783	GCCGAGGAGCAAGGTCAGAGAGACCAACCGCAGGGTGATCATCTTCAAGTGTCTTGAGAG	1842
QY	668	AlaLeuIleCysValGlySerSerYsValArgAlaArgYsAlaSerSerAspAsp	687
Db	1843	GGGTGATTGTGTGCGATGCTCCAAAGAAAGACGAGAGAGGACATCTTCAATAT	1902
QY	688	GluGlyGlyProArgThrLeuGlyYsAspSerHisArgAlaGluGlnAlaSerYsAsp	707
Db	1903	CAGAGAGGCCAAGAACACTGGAGGGGATGGCCACAGACCGAGAGAGCTAGCAAAAC	1962
QY	708	LysGlnAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGly	727
Db	1963	AAAGAAAGCA-----GATGCTCTTCTCGCAGACCCAGAAACAAGACCAACCGCACGA	2016
QY	728	SerSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGlu	747
Db	2017	AGTTCTTCAACCCGAGCCAGCGCTGGAAGGCCCTTGTGAAGGGGAGGCGCTTCCACTGGGAG	2076
QY	748	SerPheYsArgLeuValThrProArgYsYsSerYsSerYsLeuGluGlnYsAla	767
Db	2077	TCATTTAAGAGATTAGTACATCCACGAAAAAATTCAGATCAAACTGGAAGAAGAGACC	2136
QY	768	GluAspSerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTry	787
Db	2137	GAAAGACTCGGTGACAGAGAGTTGGCTTCGAGATCGAACCAAGTGAAGAGAACTTGG	2196
QY	788	ValSerIleYsYsPheIleProGlyArgArgYsYsArgAlaAspGlyYsGlnGlu	807
Db	2197	GTTTTCATTAAAGAAATTATTTCTCGAGCGCGGAGAAAGAGGCAGATGGGAACAGAA	2256
QY	808	GlnAlaThrValGluAspSerClyProValGluIleAsnGluAspAspProAsnAlaPro	827
Db	2257	CAGGCGCGCGTTTGAAGACTCGGGGCGCAGAGAGATCAATAGGACCAACCCGACGTCCCA	2316
QY	828	AlaValAlaProLeuSerGluYsAsnAlaValAlaGluArgGluYsMetGlu-----	844
Db	2317	GCTGTGTGCTCTCTGTCTGAGTACAGATCGCGTGAAGAGAGAGAACTGGAAACGACGGA	2376
QY	845	AlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValYsValSerGluGluLeu	864
Db	2377	GCTGTCGAGAAACGTAGACTGCCCACTGAAAGGGCGGCTGTGTAATGTGTGAGAGACTT	2436
QY	865	SerYsThrLeuValHisThrValSerValAlaValIleAspGlyYsThrArgAlaValThr	884
Db	2437	AGTAAGACTCTGTTCACACGTAGAGTGTGGCGGTCAATTGATGGACCAAGGGCACTGAC	2496

QY	885	SerValGIuGIuIaYserPProSerTPPIleSerAlaSerValThnGIuProLeuGIuHis	904
Db	2497	AGTCCCAAGAGGGCTCCCTTCGTGGATATCTGCTCCATGACAGAACTCTTGACAC	2556
QY	905	ThnIaGIuGIuIaMetProProValGIuGIuValThnGIuIuYAspIleIleAlaGIu	924
Db	2557	GCAAGGGAGTGGCCACACCGCTGTGGAGAGGTCACTGAAAAAGACATACCTGCAGAA	2616
QY	925	GIuThrProValLeuThnGIuThnThleuProGIuGIuYAspAlaHisAspAspMetVal	944
Db	2617	GCACTCCTGCACTCGCCGCCAGACTTTATCCAGGGGGCAAAAGATGCCATGACGATATGTC	2676
QY	945	ThnSerGIuValAspPheThnSerGIuValaValThnIaIaThnGIuThnSerGIuValaIleu	964
Db	2677	ACCAGTGAAGTGGATTTTATCTTCAGAAAGAGGACAGCGCCAGAAACCAAGAGAGCCCTC	2736
QY	965	ArgThnGIuGIuValThnGIuIaSerGIuYAlaGIuGIuThnThrAspMetValSerAla	984
Db	2737	CGCGCTGAAGACTTACCGAAGCATCAGGGGCGAAGAGACCAAGACATGATGTGTGCA	2796
QY	985	ValSerGIuLeuThnAspSerProAspThnThnGIuGIuValaThnProValGIuGIuVal	1004
Db	2797	GTITTCACAGCTGCTCCAGCTCCCGGACACACAGAGAAAGCACACCAGTTCAAGAGGTA	2856
QY	1005	GIuSerGIuValLeuAspThnGIuGIuGIuGIuIaGIuIaGIuThnGIuIaIleLeuGIuIa	1024
Db	2857	GAGGGTGGCACTGTAAGTACGAAAGAAAGAGAGCCGACAGCCAGCGCTCCCAAGCC	2916
QY	1025	ValIaAspIuYsValIuYsGIuIuSerGIuIuValProAlaThnGIuThnValGIuIaThnThr	1044
Db	2917	GTTCAGACAAAGTGAAGAAAGGACCTCCAGGTGCTTCGACCAACCAACTCTGCAGAGCA	2976
QY	1045	GIuSerIuYsAlaLeuGIuIuYsValGIuGIuValGIuGIuIaAspSerGIuValIleuAlaSer	1064
Db	2977	GGGCCGAAAGCACTGGAGAGGTGGAGAGGTGAAGAGAGATCTCGAAGTCTGGCTAAC	3036
QY	1065	GIuIuYsGIuIuYsAspValIuMetProIuYsGIuProValGIuGIuIaGIuIaGIuIuIleu	1084
Db	3037	GAGAAAGAAAGATGTGTTGGCCGGAAGGACCCGTGTCAGAAAGCTGAAACAGACATCTT	3096
QY	1085	AlaGIuGIuYsSerGIuThnGIuGIuIaIaThnProGIuIuYsLeuGIuValProGIuValThn	1104
Db	3097	GCACGGGCTCCGAGACTGTACAGGCTACCCGACAGAGCTTGAAATTCCTGAAGTCA	3156
QY	1105	AlaAspValAspHisValaIaThnCySGinValIleIuYsLeuGIuGIuIuIuMetGIuGIuIu	1124
Db	3157	GAGAGTGAAGCCCTGCCACACATCCAGGTTATCAAGACACACACACTATGGAACAG	3216
QY	1125	AlaValaIaProGIuIuYsSerGIuThnThleuThnAspSerGIuThnAsnGIuSerThPro	1144
Db	3217	GCTGTGGCCCTCGATCATCTGTAAACCTTGACAGACAGTGAACAAAGGAAGTACTCC	3276
QY	1145	LeuIaAspSerAspThnIaIaAspGIuThnGIuIuIuIuAspGIuThnIleAspSerGIuAsp	1164
Db	3277	CTCGAGATTGAGCACTCCAAACGGGACACAGCAAGAGAAAGCCGTGACAGCCAGAC	3336
QY	1165	SerIuYsAlaThnIaIaIaValaIaArgGIuIuSerGIuValThnGIuGIuGIuIaAlaIaThnAla	1184
Db	3337	AGTAAATGCCATTGGCCGCCCGTCAGACAGTCAAGGTCACTGAAGAGAGAGGCACTGCTCT	3396
QY	1185	GIuIuYsGIuGIuProSerThnThleuProAsnAsnValProAlaGIuGIuIuIuIleGIuIu	1204
Db	3397	CAGACGGAGGGGCTTCAACA---CCATCTAGTTTCCAGCCCAAGAAAGAACACAGGGA	3453
QY	1205	GIuProGIuIuYsAspValIleuGIuIuProThnGIuIuIuIuIuIuThnAlaAlaIaIaIaPro	1224
Db	3454	AAACCAAGGAAGGATGTTTCAAGAACCA---CAAGCGCTGTGCTGCGGGCACTGCTCT	3510
QY	1225	ValLeuAlaIuYsThnGIuIuValaIaGIuIuGIuGIuGIuIuValaIaAspTrpLeuAspGIuIuYs	1244
Db	3511	ATTCTGGCAAGGCTGAGGTGGGTCAAGAGGGTGAAGGCTGGCCAGTTGATGAGAAA	3570
QY	1245	ValIuYsGIu-----GIuGIuGIuValaIleValaHisSerGIuProAsnSer	1259


```
|||||:|||||
Db 3571 GTCAGACGACGACAGTGTGTTAAAGAACTGGAGGTGCTGTGCACACTGGACCCCAACAGT 3630
Qy 1260 GlnlysaAlaAaspValThrTyrAaspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
|||||:|||||
Db 3631 CAAAAGACTGCTGACTTGACACGTCGACAGTGAAGTAAATGGAAGTGGCCAGATGTCAGGAA 3690
Qy 1280 LysGluSerThrGluValGlnSerLeuSerLeuGluGluGlyGluMetGluThrAaspVal 1299
|||||:|||||:|||||:|||||
Db 3691 ACTGAGAGTAAATGAAGAACACAGATATTAGCCCGAGAAAGAGAGAGATGGGAACCGAGTT 3750
Qy 1300 GluLysGluLysAsgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThr 1319
|||||:|||||:|||||:|||||
Db 3751 GAAAGAGGAGAAACAGAGACCAACAGACGACGACCCAGTGAAGAACATGACAGCAAAACA 3810
Qy 1320 AlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAaspMetProSer 1339
|||||:|||||:|||||:|||||
Db 3811 GCTGCTCTGAGCATGAAGAAACCCACCCCTAAGCCAGTCTCTGACAGCTGACATGCCTCAC 3870
Qy 1340 SerGluArgGlyLysAlaLeuGlySerLeuGlyCysProSerLeuProAaspGlnAasp 1359
|||||:|||||:|||||:|||||
Db 3871 TCAGAGAGGGGAAAGGCACTGGCGAGCGCTTGAAGGAAAGCCCTTCTCTCCAGACCAAGAC 3930
Qy 1360 LysAlaGlyCysIleGluValGlnValGlnSerLeuAaspThrThrValThrGlnThrAla 1379
|||||:|||||:|||||:|||||
Db 3931 AAACGAGATTGCATAGAGTTCAAGTTCAAAAGCTCAGACACACCCAGTCACTCAAAACACC 3990
Qy 1380 GluAlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerProGlu 1399
|||||:|||||:|||||:|||||
Db 3991 GAAGCTGTGAAAAGGTGGAAGAACTGTGGCAACTTCAGAGATGGATGAAGTTTGGAG 4050
Qy 1400 CysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTyrThr 1419
|||||:|||||:|||||:|||||
Db 4051 TGTCCAGGTGGCGCAATCATTTACCAGCTGAGAAGCTCTCCGAAACCGGTGGCTACGGGACT 4110
Qy 1420 LeuGlnHisAlaGluAaspThrValProLeuGlyProGluSerGlnAlaGluSerIlePro 1439
|||||:|||||:|||||:|||||
Db 4111 CTTCCAGCATGAGAGGACACCGTCCCGAGGGCTGAGTCTCAGGCAGAGTCCATCCCA 4170
Qy 1440 IleIleValThrProAlaProGluSerThrLeuHisProAaspLeuGlnGlyGluIleSer 1459
|||||:|||||:|||||:|||||
Db 4171 ATAATAGTAACTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4230
Qy 1460 AlaSerGlnArgGluArgSerGluGluAaspLysProAlaGlyProAaspAlaAasp 1479
|||||:|||||:|||||:|||||
Db 4231 GCATCCCAAGAACACAGATCAGATGAAGATAACAAGCCAGATGCTGCTGCTGCTGCTGCT 4290
Qy 1480 GlyLysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeu 1499
|||||:|||||:|||||:|||||
Db 4291 GGCAAGGAGAGTGCAGCAAGAGAGAAAATCTCAGGGCTGAACCTGAGATCTTGGAACTT 4350
Qy 1500 GluSerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAaspGlnPheAla 1519
|||||:|||||:|||||:|||||
Db 4351 GAGAGTAAGACATAGATTGTCAGAGTGTCTATCAGACAGCCGTCGACCATGTTGCA 4410
Qy 1520 ArgThrGluThrAlaProGluThrHisAlaTyrAaspSerGlnThrGlnValProAla-Cy 1539
|||||:|||||:|||||:|||||
Db 4411 CGTACAGAAACAGCCCCCGAAACCCAGCTTCTGATTTACAGATCAGGTTCTCTGTGATG 4470
Qy 1539 sArgLeuAaspSerArgGluProAsnArgCysTyrThrLysMetLys----- 1554
|||||:|||||:|||||:|||||
Db 4471 CAGGCTGACAGCCAGGAGGACCAACAGATGCTGGCAAAAGATGAAGCGACCTTCAAGTC 4530
Qy 1555 -AaspAlaLysMetLysHisProValProGlnProArgGluAaspLeuValLeuThrVa 1574
|||||:|||||:|||||:|||||
Db 4531 TCCCCCAAGATGAACACTCAGTCCGTCGACCCAGGAAGACTTGGGGTTCTGTAGT 4590
Qy 1574 LeuGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVa 1594
|||||:|||||:|||||:|||||
Db 4591 TCTGAAGGCATGAGCAAGGCTTCAAGAAATGATCACCACGCTTGCAATGGAAGTGCCAGT 4650
Qy 1594 LserLys 1596
|||||
```

```
Db 4651 GTCARAG 4657
RESULT 7
LOCUS AR071282 5134 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 1 from patent US 5910442.
ACCESSION AR071282
VERSION AR071282.1 GI:7222170
KEYWORDS SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 5134)
TITLE Gelman, I.H.
JOURNAL Tumor suppressor gene
FEATURES Patent: US 5910442-A 1 08-JUN-1999;
Location/Qualifiers
1..5134
source /organism="unknown"
BASE COUNT 1511 a 1167 c 1451 g 1005 t
ORIGIN
Alignment Scores:
Pred. No.: 1 95e-206 Length: 5134
Score: 6005.00 Matches: 1232
Percent Similarity: 97.33% Conservative: 7
Best Local Similarity: 96.78% Mismatches: 27
Query Match: 74.38% Indels: 14
DB: 6 Gaps: 4
US-09-902-432-4 (1-1596) x AR071282 (1-5134)
Qy 329 GluLysThrGluProAlaSerGluGluGlnGluProAlaGluAaspThrAaspGlnAlaArg 348
Db 2 GAAAAGACAGACGACGCTCGAGGAGCAGGAGCCGCGACAGACACAGACCCAGCCAGG 61
Qy 349 LeuSerAlaAaspTyrGluLysValGluLeuProLeuGluAaspGlnValGlyAaspLeuGlu 368
Db 62 TTGTCCAGCAGACTACGAGAAGTGGAGTGCCTTTGGCAACCGAAGTGTGTATGAGAAGATGGAA 121
Qy 369 AlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAaspGluLysMetGlu 388
Db 122 GCATCTCAGAGGAGAAGTGTCTCTTTGGCAACCGAAGTGTGTATGAGAAGATGGAA 181
Qy 389 AlaHisGlnGluValValAlaGluValHisValSerThrValGluLysThrGluGluGlu 408
Db 182 GCCCACCAAGAAGTTGTCAGAGGTCCACGTGAGCACCCTGGAGAGACAGAGGAGGAG 241
Qy 409 GlnGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrGlyGluSerLeu 428
Db 242 CAGGGAGGAGGAGGAGGCTGAAGGGGGCGTGGTAGAAGGACAGGAGAAATCCTTG 301
Qy 429 ProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
Db 302 CCCCTCTGAGAAATGCTGAGCCCCCAGGAGGTCCCCCAGGAAGCTGAGCCTGCTGAGGAG 361
Qy 449 LeuMetLysSerArgGluMetCysValSerGlyLysAaspHisThrGlnLeuThrAaspLeu 468
Db 362 CTGATGAAGACAGACAGAGATGTGTCTCTGGAGGAGACCACTCACTCACTGACAGACCTA 421
Qy 469 SerProGluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMet 488
Db 422 AGTCTCTGAAGAGAAGACGCTGCCAAACACCCAGAAAGCATTTGTCAGTGGAGGATG 481
Qy 489 LeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
Db 482 CTGCTCTCTCAGGAAAGAAATCAAGGTACAGGGAAAGTCCCTTTGAAGAAACCTCTTCAGTAGC 541
Qy 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAasp 528
Db 542 TCAGGCTTAAAGAGAGTGTCTGGGAAGACAGAGAGGGGAAACGAGAGGTGGGGGAGAC 601
Qy 529 GluGluProGlyGluTyrGlnHisIleHisThrGluSerProGluSerAlaAaspGluGln 548
|||||
```

Db 602 GAAGGCGCTGGAGAAATACCAACATTCACACCAATCCCGAGAGAGTCTAGAGACG 661
Qy LysGIYgluSerSerAlaSerSerProgluGluprogluGluthrThrCysleuGlulys 568
Db 662 AAGGAGAGAGAGCTGCGCGTCCGCCGAGAGCGCTGAGAGACCAACGCTGTGGAGAA 721
Qy 569 GIYProleuGlulAlaProgluMaSpGIYgluAlaGlulGlulYThrThrSerApGIYlu 588
Db 722 GGGCCGCTGGAGACACCCAG-GATGGGAGAGCTGAGAGAGAACTACTTCGT--GGAGAG 778
Qy LysLysArgXgluGIYlIethrProTrpAlaSerPheLysLysMetValThrProLysLys 608
Db 779 AAGAAGAGAGAG--ATCACTCCCTGGGCATCTTCAAAAAGATGGTGAACCCAGAA 836
Qy ArgValAlaArgArgProSerGIuSerAspLysGIuGlulGlulYValLysSerAla 628
Db 837 CCGTC-CGAAGACCTTCTGAGAGTGAAGAAGAGAGAGCTGGAGAGAGTCAAGAGCGCC 895
Qy ThrLeuSerSerThrAspSerThrValSerGIuMetGIuMaSpGIuValLysThrValGIY 648
Db 896 ACCCTGCTCTCCACTGATAGCAAGTGTCAAAATGCAGAACTCAAACTGTGTGT 955
Qy 649 GIuGIuGIuLysProgluGluproLysArgArgValAspThrSerValSerTrpGIuAla 668
Db 956 GAGGAACAAAGCCAGAGAACCAAGCGTAGGTGATATTCTGATGTCTGGAGACA 1015
Qy LeuIleCysValGIYSerSerLysArgAlaArgLysAlaSerSerSerAspMaSpGIu 688
Db 1016 CTGATTTGTGTGGATCATCCAAAGAGAGAGAGAGAGAGCATCTTCAGATTATPAGA 1075
Qy 689 GIYGIYProArgThrleuGIY-GIYAspSerHiIaArgAlaGlulGlulAlaSerLysAsp 708
Db 1076 --GGGCCAAGACACTGGAGGGGGAGCATCAAGACAGAGAGAGCCAGCAAAACAA 1132
Qy 708 sGIuAlaGIYThrAspAlaValProIaSerThrnglulgluMaSpGIuAlaGlulYse 728
Db 1133 AAGAAGCGA-ACAGAGCGTGTCTGTCAGACCCAGAGAGAGAGACCAAGCGAAGAGAG 1191
Qy rSerSerProgluProAlaGIYSerSerSerGIuGIYgluGIYValSerThrTrpIuSe 748
Db 1192 TTCCTCACCCGAGCCAGCGGAGAGCCCTTCCAGAGGGAGAGGTCTCCACTTGGAGATC 1251
Qy 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGlulYVala-- 767
Db 1252 ATTTAAAGATTAGTCACTCAAGAAAAAATCCAAAGTCAAACTGAGAGAGAAAGA-AG 1310
Qy 768 ----GIuAspSerSerVal---GlulInLysSerThrnglulgluProSerArgGIu 785
Db 1311 CCGGAAGGACTCTAGTTGAGAGCAGGTTGTCCACTGAGATCGAACCGGTAGAGAGA 1370
Qy 785 uSerTrpValSerLysLysPheLysPheProGIYArgArgLysLysArgAlaAspGIYlu 805
Db 1371 ATCTTGCGTTTCCATTAAAGAAATTCATCCCGGACGCGGAGAGAAAAAGGCGAGATGGAA 1430
Qy 805 s-GIuGIuGIuAlaThrValGIuAspSerGIYProValGIuLysnglulMaSpPro 825
Db 1431 GGCAGAGAACACCACTGTGGAAGACTCAGGGCCAGTGAAGATTAATGAGGACGACCGCTG 1490
Qy 825 snValProIaValAlaProleuSerGIuThrAspAlaValGIuWAGGIuLysMetGIu 845
Db 1491 ATGTCCACCGCGTGTCTGTCTGAGATGATGACAGTGGAGAGAGAGAGATGGAAAG 1550
Qy 845 lAgInGIYsnThrGIuLeuProgluLeuLeuGIYValAlaValTYrValSerGIuGIuLeu 865
Db 1551 CCGAGGGGAATGGAGAGCTGCCAGCTGCTGGGGCTGTGTGA---GTGTCCGAGAGACTCA 1607
Qy 865 eLysThrLeuValHiIsthValSerValAlaValIleAspGIYThrArgAlaValThrS 885
Db 1608 GTAAAGCTCTGGTCCACACTGTGATGTGCGACATCATTATGAGGACCAAGGGCAGTCAACA 1667
Qy 885 eValGIuGIuArgSerProSerTrpIleSerAlaSerValThrGIuProleuGIuHiIst 905

Db 1668 GTGTGGAAGAGCGGCTCTCTGTGAGATATCCGCTTCCGTAAACAGAACCTTGAACACA 1727
Qy 905 hrAlaGIYgluAlaMeCProProValGIuGIuValIThrGIuLysAspLleIleAgluG 925
Db 1728 CAGCGGAGAGACCATGCGACCTGTGTAAGAGGTCACTGAAAGAACATCATTCAGAGAG 1787
Qy 925 luthrProValLeuThrGIuThrleuProgluGIYLysAspAlaHiIAspAspMetValT 945
Db 1788 AAACCTCTGTCTCAACCCAGAGCTTACAGAGGGTAAAGATGCCCATGACACATGTGTCA 1847
Qy 945 hrSerGIuValAspPheThrSerGIuAlaValIThrAlaThrGIuThrSerGIuAlaLeu 965
Db 1848 CCAAGAGAGTGGATTTCACCTCAGAGCTGTGACAGCCACAGAGACCTCAAGGCTCTCC 1907
Qy 965 rGIuGIuGIuValIThrGIuAlaSerGIYAlaGlulGIuThrThrAspMetValSerAla 985
Db 1908 GTACTGAAGAGATTACCAAGACATCGGGGGCCGAAGAGACCAAGACATGTGTCCGAG 1967
Qy 985 aLserGIuLeuThrAspSerProAspThrThrnglulgluAlaThrProValGIuGIuAla 1005
Db 1968 TTTCCAGACTGACTGACTCCCGAGACACACAGAGAGAGCCACCCAGTTGAGAGGTAG 2027
Qy 1005 luserGIYValLeuAspThrGIuGIuGIuGIuArgGIuThrGIuAlaIleleuGIuAla 1025
Db 2028 AGGCTGTGTCTAGATACAGAGAGAGAGAGCCCGCAGAGCCATCTCCAAAGCCG 2087
Qy 1025 AlaIaAspLysValIYsGIuGIuSerGIuValProAlaThrGIuThrValGIuMaSpThr 1045
Db 2088 TTGCAGACAAAGTGAAGAGAGATCCAGGTGCTGCAGAACCCAGACTGTGCAGAGAAAGG 2147
Qy 1045 lYserLysAlaLeuGIuLysValGIuGIuValGIuGIuAspSerGIuValLeuAlaSer 1065
Db 2148 GGTCAAAAGCACTGGAGAGAGGTGAGAGGTGAGAGAGAGTCCGAAGTGTGGCTTGG 2207
Qy 1065 lulYsGIuLysAspValMeCProLysGIYProValGIuGIuAlaGIYAlaGIuHiIstLeu 1085
Db 2208 AGAAAGAGAGAGACTTATGCCGAAGAGACCCGTGCGAGAGAGCTGAGACTGTG 2267
Qy 1085 lAgInGIYserGIuThrGIYgluAlaThrProgluSerLeuGIuValProGIuValIThr 1105
Db 2268 CACAGGCTGTGAGACTGAGACAGGCTACTCCAGAGAGCTTGAAGTCTCTGAAGTCAAG 2327
Qy 1105 lAspValAspHiIstValAlaThrCysGIuValIleLysLeuGIuGIuLeuMetGIuGIu 1125
Db 2328 CAGATGTAGACCAATGTCCCGAGTCCAGGTATCAACTCAGACGCTGATGGAACAGG 2387
Qy 1125 lAvalAlaProgluSerSerGIuThrleuThrAspSerGIuThrAsnGIYserThrPro 1145
Db 2388 CCGTGGCCCTGAGTCACTCCGAACCTTGACAGACAGTGAACAAATGGAAAGCACTCCCT 2447
Qy 1145 euAlaAspSerAspThrAlaAspGIYThrnglulMaSpGIuThrIleAspSerGIuAsp 1165
Db 2448 TAGCAGATTTCAGACACTGCAATGGGAGACAGACAGCAAGATGAAGAACCATGTGACCGCAGAGACA 2507
Qy 1165 eLysAlaIThrAlaAlaValArgInserGIuValIThrGIuGIuGIuAlaIaThrAla 1185
Db 2508 GTAAAGCCACTGCACTGTGAGGCGTCAAGGTCAAGAGTCAAGAAAGAGCGGCTACTGTCTC 2567
Qy 1185 lulYsGIuGIuProSerThrleuProMaAsnValProAlaGIuGIuGIuHiIstGIYGIu 1205
Db 2568 AGAAAGAGAGGCTTCCAGACATCACTAATATGTTCCAGCCAGAGAAACATGGGGAAG 2627
Qy 1205 luproGIYArgAspValleuGIuProThrnglulgluLeuThrAlaAlaValPro 1225
Db 2628 AACAGAGAGAGATGTCTTGAACCTACAGACAGAGAGACTGTGCTGCAAGCGGCGG 2687
Qy 1225 AlLeuAlaLysThrGIuValaGIYnglulGIuGIuValaAspTrpleuAspGIYGIuLys 1245
Db 2688 TCTGGCAAAAGACTGAGGTGGGTCAAGAGGGTGAAGTGTGACTGTGTGATGGAGAAAAAG 2747
Qy 1245 AlLysGIuGIuGIuGIuValaPheValHiIstSerGIYProAsnSerGIuLysAlaIaAsp 1265
Db 2748 TCAAAGAGAGAGAGAGGTGTTGTACCTTGAACCCAGACAGTCAAAAGGCTGCTGATG 2807

QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGluLysAsnGly 56
 Db 312 ACCGGGAGCCCGGCGATGCTGCTCGGAGCCCGGACCAAGACTCTCAGAAAGATGCT 371
 QY 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluLysAsn 76
 Db 372 CAGCGTCACCATCAATGAGCTAGCTAGCAAGATGAGCTCAGCTCCAGAGAGGTGAC 431
 QY 77 GlnGluGlyGln-----GlnGlu 82
 Db 432 CTAATATGGCCAGAAAGAGAGCCCTGAACGCTCAGAGAGCCCTAAACAGCCAGAGAGAGAA 491
 QY 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
 Db 492 GAAGTCATTTGTCACAGAGGTTGAGCAGAGAGACTCTGAAGATGTGAGCGAAAGAGACTCC 551
 QY 103 ValGluGluMetAlaAlaAsnSerThrAlaValAlaGluAspIleThrLysAspGlyGlnGlu 122
 Db 552 GATAAAGAGATGGCTAATGAAGTCAAGCGTTGTTCAGACATCAGATGATGGGAGAGAG 611
 QY 123 GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetVal 142
 Db 612 GAGACACCCGAAATATATGACACAGATCTCTTCAGAAAGCAATTTAGAGAGACTAACA 671
 QY 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
 Db 672 CAACCCATCGAGTCCAGGCTAATGATATGATTTAAGAGAGTGTAAAGTTGTGTGGC 731
 QY 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
 Db 732 TTTAAATTCACGTGAAAAAGATTAAGACAGAGAAAGCTGACACGTCCAGCTACTACT 791
 QY 183 ValLysLysAspGluGlyGluGluLysAlaGluLysValGlyAlaGlyAspHisGlnGlu 202
 Db 792 GTGAAGAAAGATGAAGGGGAGGAGCA-----GCGGGGGCTGGGCAACACAGAGAC 842
 QY 203 ProSerValGluThrAlaValAlaGlyLysLysLysLysLysLysLysLysLysLysLys 222
 Db 843 CCCACCTT-----GGGGCTGGAAGAGCAGCATCCAAAGAAAGCAACCAACATCT 896
 QY 223 ThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIleProLeuGln 242
 Db 897 ACAGAGAAACCCGAGAGAGACCTTGAAAGCTGAGCAAAAGCCAGCAAGAAATTTCTCCCA 956
 QY 243 AlaGluSerAspGlnAlaAlaGluGluGluLysAspGluGlyGluGluLysGlnGlu 262
 Db 957 GCCGATCTGGCCAGAGCAAGT---GAGGATGCAAAAGAGAGAGAGAGAGAGAGAGAA 1013
 QY 263 LysGluProThrLysSerProGluSerProSerProValAsnSerGluThrThrSer 282
 Db 1014 AAAGAACCTGACAGATCTGCAATCTCCGACTAGTCCCGTGAACAGTGAACAGATCA 1073
 QY 283 SerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSerPheLys 302
 Db 1074 ACCCTCAAAAATTTCTCACTCAAGGTTGGCGCGGTGCGCAAAAAGCCAGTTTCAAG 1133
 QY 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluLysGluLys 322
 Db 1134 AAGCCGAAAGAGATGAGTGAAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1193
 QY 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
 Db 1194 GTAGACACAG 1253
 QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspLys 353
 Db 1254 CAAGCCCAACCCAGAGAGCGGACAGAAAGTGCACAGAGAGAGAGAGAGAGAGAGAGAG 1313
 QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluLysSerGluGlu 373
 Db 1314 GAGAAAGTTGAGCTGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1373

QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
 Db 1374 AAACCTGCTCCGTGGGAGACAGAAAGTGTGTATGAGAAATAGAAAGTCCACCAAGAGAG 1433
 QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGluGluGlyGly 412
 Db 1434 GTTGTGGCCGAAAGTCCACGTACGACCCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1484
 QY 413 GlyLysAlaGluGlyGlyValValValGluGlyThrGlyLysSerLeuProGluLys 432
 Db 1485 -----AAACGAGGTGAGAGAAACAGAGAGGTCTGTGTCCAGCTGAAGAA 1529
 QY 433 LeuAlaGluProGlnGluValAlaProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
 Db 1530 TTGTTGAATATGATGACAGAACTTCAGAGAACCCAGACCTGCAAGAGAGCTGTAAAGCTC 1589
 QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
 Db 1590 AAAGAACGTGTGTTCGAGAGAGACCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1649
 QY 473 LysThrLeuProLysHisProGluGlyLysValSerGluValGluMetLeuSerSerGln 492
 Db 1650 AAGTGTCTGCCAAACCCCGGAGAGGCTGTGAGTGAAGTGAAGATGCTGTATCAAG 1709
 QY 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLysLys 512
 Db 1710 GAGAGAAATGAAGGTGCGAGGAGAGTCCATTAAGAAAGCTTTTACAGACATCGGCTTAA 1769
 QY 513 LysLeuSerGlyLysLysGluLysGlyLysArgGlyGlyGlyLysArgGluGluProGly 532
 Db 1770 AAGCTTTTGAGAAAGAAACAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1826
 QY 533 GluTyrglnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyLysSer 552
 Db 1827 GAGACACACTCAGTTCACAGCGGATTTCTCGGACAGCCAGAGAGAGAGAGAGAGAGAGAG 1886
 QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
 Db 1887 TCTGCTCATCTCCCTGAGAGACCCGAGAGATCACTGTCTGAGAAAGGGCTTAGCCAG 1946
 QY 573 AlaProGluAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
 Db 1947 GTGCAGCGAGAGTGGGAGAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 2006
 QY 593 GlyIleThrProTTPAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
 Db 2007 GGTCATCTCCCTGGGATCTTCAAAAGATGGTGAAGCCCAAGAGAGAGAGAGAGAGAGAG 2066
 QY 613 ProSerGluSerAspLysGluGluGluLysValLysSerAlaThrLeuSerSer 632
 Db 2067 CCTTCGAAAGTGAATAAAGAGATGAGCTGAGCAAGGTCAAGAGAGAGAGAGAGAGAGAG 2126
 QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
 Db 2127 ACCGAGACACAGCCTCTGAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 2186
 QY 653 ProGluGluProLysArgArgValAspThrSerValSerTTPGluAlaLeuIleCysVal 672
 Db 2187 CCGAGAGAACCAAGCCGAG 2246
 QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyProArg 692
 Db 2247 GGATCATCAAGAAAG 2306
 QY 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluLysSerLysAspLysGluAlaGlyThr 712
 Db 2307 GCATATGAG 2366
 QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGlu 732
 Db 2367 GACGGGATCTCTTGCTGTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2426
 QY 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTTPGluSerPheLysArgLeu 752

Db 2427 CAAGCTGGAAGCCCTCCGAAGGGAGGGCGTTTCCACCTGGGAGTCATTAAAGGTTA 2486
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp----- 769
Db 2487 GTACGCGCAAGAAAAAATCAAGCTCAAGCTGGAAGAGAAAGCGAAGACTCCATAGCT 2546
Qy 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal 788
Db 2547 GGGTCTGGTGTAGAACATTCCACTCCAGACACTGAACCCGGTAAAGAAATCTGGGTC 2606
Qy 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
Db 2607 TCAATCAAGAAAGTTTATCTCTGGACGAAGGAAGAAAGGCCAGATGGGAACAAGAACAA 2666
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2667 GCCCCTGTTGAAGACGCGAGGCCCAACAGGGGCCAACGAAGATGACTCTGTATGTCCTCGGCC 2726
Qy 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2727 GTGGTCCCTCTGTCTGATGATGCTGTAGAAAGGGAGAAAAATGAGGCACACAAGCC 2786
Qy 846 GlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2787 CAAAAAGCGCAGAGCAGCCCGCAGAGAGCCACTGAGGTGTCCAGAGAGCTCAGC 2846
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2847 GAGAGTCAGGTTCATATGATGCGCAGCAGCTGCTGACGGGACGAGGCAGCTACCAT 2906
Qy 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2907 ATTGAAGAAAGGTCTCCTCTCTGTGATATCTGCTTCAGTGACAGAACCTCTTGAACAAGTA 2966
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
Db 2967 GAAGCTGAAGCCGCACCTGTTAACTGAGGAGGTATTGAAAGAGAAAGTAATTGCAGAA 3026
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3027 GAACCCCCACCGTTACTGAACCTCTGCCAGAGAACAGAGAGGCCCGCGGCGCACCGTC 3086
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3087 GTTAGTGAGGCGGAATTGACCCCGAAGCATCTGCTGTAAGAGACCCAGAAATGGTGTACGA 3146
Qy 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluThrThrAspMetValSerAla 984
Db 3147 GGTTCGAGAGAAAGAACCGAAGCATCTGCTGTAAGAGACCCAGAAATGGTGTACGA 3206
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3207 GTCTCCAGTTAACCGACTCCCGACAGACACAGAGAGAGGCCACTCCGGTGCAGGAGTG 3266
Qy 1005 GluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3267 GAAGGTGGCGTACCTGACATAGAGAGAGAGAGAGAGGGGACTCAAGAGGTCTCTCAGGCA 3326
Qy 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3327 GTGCAGAAAAGTGAAGAGGAATCCAGCTGCTCGCACCCCGTGGCCGAGAGATGTG 3386
Qy 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluValGluGlu 1057
Db 3387 CTTTCAGCTGTGCAGAGCA-----GAGCGAGAAAGACCCAGAAAGAG 3428
Qy 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
Db 3429 CAGGCTGAAGCGTCGGGTCTGAAGAAAGAGAGCGATGTAGTTGTAAGATAGATGCTCAG 3488
Qy 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097

Db 3489 GAGGCAAAAAAGCTGAGCCTTTTACACAAGGAAGGTGGTGGGGCAGACCACCCCGAAGAGC 3548
Qy 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3549 TTTGAAAAAGCTCCTCAAGTCACAGAGAGCATAGAGTCCAGTCCAGCTTGTAAACCACTGT 3608
Qy 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db 3609 CAAGCCGAAACCTTAGCTGGGTAAATAATCAAGAGATGGTGTATGGAACAGGCTATCCCC 3668
Qy 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3669 CTTGACTCGGTGGAACCCCTCAGACAGTGAGACTGTGAAGACCCCGCTGATGCGAC 3728
Qy 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
Db 3729 TTTGACGACACAGGCAACCCAGAAAGACGAGATTGTGGAATCCATGAGGAGAATGAG 3788
Qy 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
Db 3789 GTCCGATCTGTGTACCCAGTCAGGGGCGCACAGAAGCAGAGGACGTTCTGTCACAGAAAG 3848
Qy 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly 1207
Db 3849 AGGCTCCAGCA---CCTTCAGTTTTGTCTCCAGGAAGAACTAAAGAACAACTCAAG 3905
Qy 1208 ---ArgAspValLeuGluProThrGlnGlnLeuThrAlaAlaValProValLeu 1226
Db 3906 ATGGAAGACACTCTAGAGCATACAGATAAAGAGGTCTCAGTGGAACCTGTATCCATCTG 3965
Qy 1227 AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys 1246
Db 3966 TCAAAGACTGAGGGGACTCAA-----GAGGCTGACAGTATGCTGATGAGAAACCAAA 4019
Qy 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4020 GAGCTACCATTTTTCAGAGGACTTGAGGGGTCTATAGACACAGGCATACACGTCAGTCGG 4079
Qy 1260 GlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4080 GAAAGGTCACTCAAGTTGCCCTTAAAGGTGAAGGGACAGAGAAAGTGAATGTAAAAAG 4139
Qy 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
Db 4140 GATGATGCTCTTCAACTGCAGAGTCACGCTAAGTCTCTCCATCCCTCCCGTGGAGAGAG 4199
Qy 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
Db 4200 ATGGTAGTTCAAGTCAAGGGGAGAAAACAGAAAGCAGAGCCAAACCCATGTGAATGAAG 4259
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeu 1333
Db 4260 AAGCTTGACACCAACAGCTGTTACCGTATCTGAAGAGGTCTAGTAAGCAGCTCCTCCAG 4319
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
Db 4320 ACAGTGAATGTGCCCATCATAGATGGGCAAGAAAGTCAAGAGTGTGGAAGGAAGCCCT 4379
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4380 CCTCCCTGCTAGGTCAAGAGGAGGAGTATGCACCAAAATTCAGATTCAGAGCTCTGAG 4439
Qy 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
Db 4440 GCATCATCTCTACAGCGGCTGCAGAGGAGAAAGGTCTTAGAGGAAATGCCAAC 4499
Qy 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4500 ATTTAGAAACAGGTGAAAGCTTGGAGCCTGCAGGTGCACATTTAGTTCTTGGGAAGAGAA 4559
Qy 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4560 TCCTCTGAAAAAAATGAAGACTTTTGGCGCTCATCCAGGGGAGAGTGTGTGCTCCACAGGG 4619

Qy 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrIleu 1450
 Db 4620 CCCCAGCTGTACGCAAAATCGACACAGATGATATCTACTACCAAGAAAGCTTA 4679
 Qy 1451 HisProAlaGluGlnGluIleSerAlaSerGlnArgGluArgSerGluGluIleu 1470
 Db 4660 AGTTCGACCTGGAGAGAAACCATCATCTAGATGTGAGTGAATGATCGAT 4739
 Qy 1471 LysProAlaGluIleProAlaAlaAspGlyLysGluSerThrAlaIleGluLysValIleu 1490
 Db 4740 GAGCAGGTTCCTTGCACAGAGTC-----AAGAGAGTGTGACATTGAGAGATTTTA 4793
 Qy 1491 LysAlaGluProGluIleIleGluIleuGluSerLysSerAsnLysIleValIleuAsnVal 1510
 Db 4794 GAGCCTGAAAAATGGATTTTGGACTTGAGACCAAAAGCAGTAACTTCTCCAAAACATC 4853
 Qy 1511 IleGlnThrAlaValAlaAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
 Db 4854 ATCCAGACAGCCGTTGACCGATTGTACGTACAGAAAGAACCCAGCAATGTTGACG 4913
 Qy 1530 TyrAspSerGlnThrGlnValProAlaCysArgLysAspSerArgGluProAlaArgCys 1549
 Db 4914 TCTGAGTTACAGACACAGCTCAGCTGATAAAGCTGACAGCAGGACCTGACAGAA 4973
 Qy 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProValProGlu 1564
 Db 4974 ACGGAAAGAAAGAGAGAGACCTCAGGCTCTGCACAGATGAAACCAATTAATTCTCA 5033
 Qy 1564 nProArgLys-----AspLeuGlnValIleuThrValLeuGluAla 1577
 Db 5034 GCCAAGAGAGACTCAGAGTCAACCGCATGTGACAAAGCAATTCGTGATTTTCCAAAGAC 5093
 Qy 1577 atTrpAlaGlnProArgLysCysLeuProArgLeuGlnIleuLysAlaProVal 1594
 Db 5094 ATGATGAGAGCTCAGAAAGAACATGACTGTGAGGTGAAGGTTCCACTG 5145

RESULT 9 AX329766 6608 bp DNA linear PAT 09-JAN-2002
 LOCUS AX329766
 DEFINITION Sequence 275 from Patent WO0194629.
 ACCESSION AX329766
 VERSION AX329766.1 GI:18102744
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horvath, S., Soppet, D.R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets
 Patent: WO 0194629-A 275 13-DEC-2001;
 JOURNAL Avalon Pharmaceuticals (US)
 FEATURES
 source 1..6608
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 2091 a 1418 c 1827 g 1272 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,85e-145 Length: 6608
 Score: 4300.50 Matches: 967
 Percent Similarity: 68.81% Conservative: 189
 Best Local Similarity: 57.56% Mismatches: 408
 Query Match: 53.27% Indels: 116
 DB: 6 Gaps: 36

US-09-902-432-4 (1-1596) x AX329766 (1-6608)

Qy 1 MetGlyAlaGlySerThrGlnArgSerProGluGln--ProAlaGlySerAsp 19

Db 192 ATGGCCCGGAGACTCCACGAGCAGCGCAGCCGAGCGCCGAGGGAGCTCC 251
 Qy 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer--GlyAla 38
 Db 252 ACGCCGGGTGACCCGAGCCAGCCAGCGCGCGGCCCTTCGCGCGAGCGGCCGACAGACC 311
 Qy 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
 Db 312 ACCGGGACCCCGCATCGCTGCTCGGACCCGCGCACCAAGCTCTACAAAGAAATGGT 371
 Qy 57 GluLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluLysn 76
 Db 372 CAGCTGCCACCATCATATGCTGCTAGCTGAGCAATGAGCTCAAGCTTCAGGAGGATAC 431
 Qy 77 GlnGluGlnGln-----GluGlu 82
 Db 432 CTAAATGGCCAGAAAGAGAGCCCTGAACGGTCAAGGAGCCCTTAAACAGCCAGAGAGAGAA 491
 Qy 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
 Db 492 GAAGTCATTTGACAGGAGGTGACAGAGACTCTGAAGATGTGAGCGAAAGAGACTCC 551
 Qy 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu 122
 Db 552 GATTAAGAGATGGCTACTTAAGTCAAGCGGTGTTCACGACATCAAGATGATGGCAGAG 611
 Qy 123 GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluIleuVal 142
 Db 612 GAGAACCGAAT--ATGAAACAGATTCCTTCAGAAACCAATTTGAAGAGCTAAC 668
 Qy 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
 Db 669 CAACCCAGTGTCCAGCTCAATGATATTGAAAGGGGTTAATTTGTTGTC 728
 Qy 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGluLeuLeuThr 182
 Db 729 TTTAATTCAGCTGTAAGAAAGATAGACAGAGAGCTGACACTGTCACTACTCTACT 788
 Qy 183 ValLysLysAspGluGluGluValAlaGluAlaLysValGlyAlaGlyAspHisGlnGlu 202
 Db 789 GTGAAGAAAGATGAAGGGGAGAGCA-----GAGGGGGCTGGCGACACAGAGAC 839
 Qy 203 ProSerValGluThrAlaValAlaGlyLysSerLysSerLysGluLeuLysGlnSer 222
 Db 840 CCCAGCTT-----GGGGCTGGAGAGACAGCATCCAAAGAAAGGAAACCAAAATCT 893
 Qy 223 ThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIleProLeuGln 242
 Db 894 ACGAGAAACCCGAGAGAGACCTTGAGCGTGAAGCAAGCCAGCAAGAAATTTCTCCCA 953
 Qy 243 AlaGluSerAspGlnAlaAlaGluGluLysAlaLysAspGluGluGluLysGlnGlu 262
 Db 954 GCCGATCTGGCCAGACAGTG--GAGGAATGCAGAAAGAGAGAGAGAGAAACAGAA 1010
 Qy 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
 Db 1011 AAAGAACTAGCAAGCTGCAAGATCTCCGACTAGTCCCGTGAACCAAGAGAAACAGAGTCA 1070
 Qy 283 SerPheLysLysPhePheThrHisGlyThrAlaGlyTrpArgLysLysThrSerPheLys 302
 Db 1071 ACCTTCAAAAATTTCTTCACTCAAGGTGGCGGCTGGCGCAAAAACCAAGTTTCAGG 1130
 Qy 303 LysSerLysGluAspAspLeuGlnThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
 Db 1131 AAGCCGAGAGAGATGAGTGAAGCTTCAGAGAGAAAGAAAGAGAGAGAGAGAGAGAGAG 1190
 Qy 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGln----- 337
 Db 1191 GTAGACACAGAAAG 1250
 Qy 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerLysAspThr 353

Db 1251 CAAGCCACACAGGAGCGGCAGAAAGTGCCACAGAGCCCGGTTATCAGCTGAATAT 1310
Qy 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1311 GAGAAAGTTGAGCTGCGCTCAGAGAGCAAGTCTAGTGGCTCGCAGGAGACCTTCTGAAGAG 1370
Qy 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
Db 1371 AAACCTGCTCGTTGGCGACAGAAAGTGTGTATGAGAAATAGAGTCCACCAAGAAAGAG 1430
Qy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlnGlyGly 412
Db 1431 GTTGTGCGCAAGTCCACGTCAGCACCGTGGAGGAGAGAACCGAAGAGCAG----- 1481
Qy 413 GlyGluAlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProProGluLys 432
Db 1482 -----AAAACGGAGGTGGAAGAACACAGCAGGGTCTGTGCCAGCTGAAGAA 1526
Qy 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
Db 1527 TTGTTGGAATGAGTGCAGAACCTCAGAAAGCCGAACCTGCCAAGGAGCTGGTGAAGCTC 1586
Qy 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1587 AAAGAAACGTTGTTCCGAGAGGACCTTACACAGGAGCTGACCTCAGTCTGTATGAG 1646
Qy 473 LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGln 492
Db 1647 AAGTGTGTCCTCAACCCCGAAGCGGTGTGAGTGAGGTGGAATGCTGTATCATCAG 1706
Qy 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerGlyLeuLys 512
Db 1707 GAGAGAAATGAAGTGCAGGAAGTCCACTAAAGAGCTTTTACCAGCACTGGCTTAAAA 1766
Qy 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyAspGluGluProGly 532
Db 1767 AAGCTTTCGGAAGAACACAGAAAGGANAAGA---GGAGGAGAGACGAGGATCAGGG 1823
Qy 533 GluThrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1824 GAGCACACTCAGGTTCAGCGCATTTCTCCGACAGCAGGAGGACAAAGGCGGAGAGC 1883
Qy 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1884 TCGCTCATCCCCCTGAGGAGCCGAGGAGATCAGTGTCTGGAAGAGGCTTAGCCGAG 1943
Qy 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1944 GTCAGCAGGATGGGAAGCTGAAGAGGAGCTACTTCCGATGGAGAGAAAGAGAA 2003
Qy 593 GlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArg 612
Db 2004 GGTGTCACTCCCTGGGCATCATTCAAAGATGTGTGACGCCCAAGAACGCTGTAGACGG 2063
Qy 613 ProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
Db 2064 CCTTCGAAAGTATAAAGAGATGAGCTGGACAAGGTCAAGAGCGCTACCTTGTCTTCC 2123
Qy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
Db 2124 ACCGAGAGCAGACCTCTGAATGCAAGAGAAATGAAAGGAGCGCTGGAAAGGCCAAAG 2183
Qy 653 ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
Db 2184 CCGAAGAACCAAGCCAAAGTGGATACCTCAGTATCTTTGGGAAGCTTTAATTTGTGTG 2243
Qy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg 692
Db 2244 GGATCATCCCAAGAAAGAGCAAGGAGGTCTCTTCTGATGAGGAAGGGGACCAAAA 2303
Qy 693 ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2304 GCATGGGAGAGACCCACCAAGAGCTGTAGCGCCGGAAGAACAAAGAGACGGGACA 2363

Qy 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGlu 732
Db 2364 GACGGGATCCTGCTGTTCCCAAGAACATGATCCAGGGCAGGGAAGTCTCTCCCGGAG 2423
Qy 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeu 752
Db 2424 CAAGCTGGAAAGCCCTACCAAGGGGAGGCGTTTCCACCTGGGAGTCATTTAAAGGTTA 2483
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp----- 769
Db 2484 GTCACGCCCAAGAAAAAATCAAGTCCAAGCTGGAAGAGAAAAAGCAGACTCCATAGCT 2543
Qy 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal 788
Db 2544 GGTCTCGTGGTGTAGAACATTTCCACTCCAGACACTGAACCCCGTAAAGAGAATCTCGGTC 2603
Qy 789 SerIleLysLysPheIleProGlyArgArgLysLysLysArgAlaAspGlyLysGlnGluGln 808
Db 2604 TCAATCAAGAAAGTTTATCTCTGGACGAAGAGAAAGGCCAGATGGGAAACAAAGAACAA 2663
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2664 GCCCTGTTGAAGACCGCAGGCCCAACAGGGGCCAACGAAGATGACTCTGATGTCCTCGGCC 2723
Qy 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2724 GTGGTCCCTCTGTCTGAGTATGCTGTAGAAAGGAGAGAAATGGAGCAGCAAGGCC 2783
Qy 846 GlnGlyAsnThrClnLeuProGlnLeuLeuGlyAlaValTyValSerGluGluLeuSer 865
Db 2784 CAAAAAGGCGCAGAGAGCCCGCAGAGAGAGGAGGAGGCTGAGGTGTCCAAAGGAGCTCAGC 2843
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2844 GAGAGTCAGGTTTCATATGATGGCAGCAGCTGTCTGACGGGACGAGGCGAGCTACCAT 2903
Qy 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2904 ATTGAAGAAAGGTCTCCTCTTGGATATCTGCTTCACTGACAGAACCTCTTGAAACAAGTA 2963
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
Db 2964 GAAGCTGAAGCCGCACTGTAACTGAGGAGGTATTGGAAGAGAGAGTAATTCAGAAGAA 3023
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3024 GAACCCCCACCGTTACTGAACTCTGCCAGAGACAGAGAGGCGCCGGGCGACACGGTC 3083
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3084 GTTAGTGAGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGGCCATTG 3143
Qy 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3144 GGTTCGAAGAGAGAACCCGAAGCATCTGCTGCTGAAGAGACCAACAGAAATGTTGTGCA 3203
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3204 GTCTCCAGTTTAAACGACTTCCCGACACACAGAGGAGGCCACTCCGGTGCAGGAGGTG 3263
Qy 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3264 GAAGGTGGCGTACCTGACATAGAGAGCAAGAGAGGCGGACTCAAGAGGTCTCTCCAGGCA 3323
Qy 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3324 GTGGCAGAAAAAGTGAAGAGGAATCCAGCTGCCCTGGCAGCCGCGTGGCCAGAGAATGTG 3383
Qy 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3384 CTTTACGCTGTGCAGAGACA-----GAGCAGAAAGACCAAGAGAG 3425

Db 2244 GGATATCCAAAGAAAGCAAGAGAGCTCTCTTCATGATGAGGAAGGGGACCAAA 2303
 QY ThrLeuGlyIAspSerHisArgAlaGluIValAspLysAspIleGluIThr 712
 Db 2304 GCAATGGGAGAGACCAACAGAAAGCTGATGAGCCGGAAGAAAGACAAAGAGCGGGACA 2363
 QY AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerProGlu 732
 Db 2364 GACGGGATCTCTGCTGTTCCCAAGAACATGATCCAGGCGAGGAAATCTCTCCGGGAG 2423
 QY ProAlaGlySerProSerGlnGluGluIValSerThrTriPgluSerPheLysArgLeu 752
 Db 2424 CAAGCTGGAAGCCCTTACCAAGAGGAGGGCCCTTCCACTGGGAGCTTAAAGGTTA 2483
 QY ValThrProArgLysLysSerLysSerLysLeuGluLysAlaGluAsp 769
 Db 2484 GTACGCCCAAGAAAAATCAAAAGTCCAAAGCTGAAGAGAAAGCGAAGACTCCATAGCT 2543
 QY ---SerSerValGluGluIleuSerThrGluIleGluProSerArgGluIleuSerThrVal 788
 Db 2544 GGATCGGTGTAGAACATTCACCTCCAGCACTGAACCCGGTAAAGAAATCTCGGGTC 2603
 QY SerIleLysPheIleProGlyArgLysLysArgAlaAspGlyLysGlnGluGln 808
 Db 2604 TCAATCAAGAGATTATCTCTGACCAAGAGAAAGAAAGCCAGATGGGAAACAGAACAA 2663
 QY AlaThrValAlaAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
 Db 2664 GCCCGTGTGAAGAGCAGAGGCGCAACAGGGGCGCAAGAAATGACTGTGATGCTCCGGCC 2723
 QY ValValProLeuSerGluIleValAsnAlaValAlaGluIleLysMetLys 845
 Db 2724 GTGGTCCCTCTGTGATGATGATGCTGTAGAAAGGAGAAATGAGAGCAGCAAGAGCC 2783
 QY GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrosValSerGluIleuSer 865
 Db 2784 CAAAAAGGCGAGAGACGCCGAGCGAAGAGCCACTGAGGTGTCCAAAGAGCTCAGC 2843
 QY LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
 Db 2844 GAGAGTCAGATTATGATGATGAGCAGCTGTGCTGAGCGAGCAGAGGAGCTACATT 2903
 QY ValGluGluIleuSerProSerThrIleSerAlaSerValThrGluProLeuGluIleuThr 905
 Db 2904 ATTGAAGAAAGGCTCTCTCTGATATCTGCTTCACTGACAGAACCTCTTGAACAAGTA 2963
 QY AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla 924
 Db 2964 GAAGCTGAAGCCCACTGTTAACTGAGAGGTATTTGAAAGAGAAATTTGCAAGAA 3023
 QY GlnThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
 Db 3024 GAACCCCAACGGTTACTGAACCTCTGCCAGAACAGAGAGGCCCGGGGCGACAGCTC 3083
 QY ThrSerGluValAspPheThrSerGluAlaValThrAlaThrIleuThrSerGluIleu 964
 Db 3084 GTTAGTGAGCGGGAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCGAGGCCATTG 3143
 QY ArgThrGluGluValThrGluIleAspGlyAlaGluGluIleuThrAspMetValSerAla 984
 Db 3144 GGTTCGGAAGAGAACCGGAGCATCTGCTGTGAAGAGACCAAGAAATGTGTGACGA 3203
 QY ValSerGlnLeuThrAspSerProAspThrThrGluGluIleuValThrProValGlnIleu 1004
 Db 3204 GTTCCCAAGTTAACGACTCCCAAGACACAGAGAGAGGCCACTCGGTGCGAGAGGTG 3263
 QY GluSerGlyValLeuAspThrGluGluGluIleuArgGlnThrGlnAlaIleLeuGlnAla 1024
 Db 3264 GAAGGTGGCTGACTGATGATGAGAGCAAGAGAGCGGACTCAAGAGGTCTCCAGGCA 3323
 QY ValAlaAspLysValLysGluIleuSerGlnValProAlaThr 1038
 Db 3324 GTGGCAAGAAAGTGAAGAGATCCAGCTGCTGGCACCGGTGGGCCAGAAAGATGTG 3383

QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluIleu 1057
 Db 3384 CTTCAGCTGTGGCAGAGAGCA 3425
 QY AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
 Db 3426 CAGGTGAAGGCTGGGCTGAAGAGAGAGCGATGTAGTGTGAAAGTGAATGATGCTCAG 3485
 QY GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
 Db 3486 GAGGCAAAAACGAGCCCTTTTACACAAGGAGAAAGGTGGGGGAGACCAACCCAGAAAGC 3545
 QY LeuGlu---ValProGluValThrAlaAspValAspHis 1113
 Db 3546 TTTGAAAGGCTCTCTCAAGTACAGAGACGTAAAGTCACTGAGCCTGTATCAACCTTGT 3605
 QY GlnVal 1114
 Db 3606 CAAGCCGAAACCTTAGCTGGGCTAAATCAACAGAGATGTGATGAGACAGGCTATCCCC 3665
 QY ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
 Db 3666 CCGTACTCGGTGAGAAACCCCTACAGACAGTGAAGTGAAGAGACCCCGTACCGGAC 3725
 QY SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
 Db 3726 TTTGACGACACAGGACCAACCCAGAAAGCAGATTTGGAAATCCATGAGAGAGATGAG 3785
 QY ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaIleThrAlaGluLysGlu 1187
 Db 3786 GTCCGATCTGTGATCCCACTACAGGGGCAACAAGAGAGAGAGGCTCTCCACAGAAAGAG 3845
 QY GluProSerThrLeuProAsnAsnValProAlaGlnGluGluIleuIleGluIleuProGly 1207
 Db 3846 AGGCTCCAGCA---CTTCCAGTTTGTGTTCAGAGAAAGAACTAAAGAAATCAAAAG 3902
 QY 1208 ---ArgAspValLeuGluProThrGlnGlnGluIleuThrAlaAlaValProValLeu 1226
 Db 3903 ATGGAAGCACCTTGAACATACAGATTAAGAGGTGACGTGAGAAATCTGTCCATTGTG 3962
 QY 1227 AlaLysThrGluValGlyGlnGluGlyValAspTrpLeuAspGlyGluLysValLys 1246
 Db 3963 TCAAAAGCTGAGGGGACTCAA-----GAGGCTGACCACTTCTGATGAGAAACCAAA 4016
 QY 1247 -----GluGlnGlnIleuValPheValHisSerGly-----ProAsnSer 1259
 Db 4017 GACGTACCAATTTTTCGAAAGGACCTGAGGGGTCTATAGACACAGGCAATACAGTCACTGG 4076
 QY 1260 GlnLysAlaAlaAspValThrTyrosSerGluValMetGlyValAlaGlyCysGlnGlu 1279
 Db 4077 GAAAAGGTCACTGAAAGTTCCTTAAAGTGAAGGAGGAGCAAGAAAGCTGAATGTAAAAAG 4136
 QY 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluIleu 1294
 Db 4137 GATGATGCTCTTGAACCTGACAGTCAACCTTAAGTCTCTCAATCCCGGTGAGAGAGAG 4196
 QY 1295 MetGlnThrAspValAlaGluLysGluLysArgGluThrLysProGluGlnIleuValSerGluGlu 1314
 Db 4197 ATGTGATGTTCAAGTGAAGAGGAGAGAAACAAGAGAGACCAACCATGTGAATGAAGAG 4256
 QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrosLysProValLeu 1333
 Db 4257 AAGCTTGAAGCAGAAACAGCTGTACCTGATCTGAAGAGGTCAAGTCAAGCTCTCCAG 4316
 QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
 Db 4317 ACAGTGAATGTGCCATATGATGAGGCGAAAGAGAGTACAGAGTTTGAAAGAGAGCCCT 4376
 QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnIleuSerLeuAsp 1372
 Db 4377 CTTCCCTGCTGATGATCAAGAGAGGAGGAGGATGACCAAAATTTCAAGTTCAAGGCTGTAG 4436

```
QY 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
Db 4437 GCATCTACTTCAACAGCGCTGCAGAGAGGAAAGCTTCTAGAGAAATGCCAAC 4496
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4497 ATTTAGAACAGGTGAACGTTGAGCGCTGCAGGTGCACATTTAGTTCTGGAAGAAA 4556
QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4557 TCCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGAAGATGCTGTGCCACAGGG 4616
QY 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
Db 4617 CCCAGTGTGAGGCAAAATGACACCATGATAGTATCTGCTACTACCAAGAAAGGCTTA 4676
QY 1451 HisProAspLeuGlnGlyIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
Db 4677 AGTTCACCTCGAAGAGAGAGAAACCATCATCTGAAGTGAAGTCAGATGAGTGCAT 4736
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4737 GAGCAGGTGCTTCCAGGAGGTC-----AAAGTGAGTGAGCAATTGAGGATTTA-- 4787
QY 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4788 -----GAGCTGAAATGGGATTTTGGAACTTGAGACCAAGCAAAAGCAGTAACTTGTCAA 4841
QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4842 AACATCATCAGACAGCGCTTGACCACTTGTGCTGACAGAGAAACAGCCACCGAAATG 4901
QY 1528 HisAlaTyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4902 TTGACGCTCAGTTTACAGACACAGCTCAGCTGATAAAGCTGACAGCCAGGACGCTGA 4961
QY 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa 1562
Db 4962 CAGGAACCGAGAAAGAGGAGGAGAACTCAGGCCCTCTGCACAGGATGAACACCAAT 5021
QY 1562 lProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5022 ACTTCACCAAGAGGAGTACAGTCAACCGCAGTGGGACAGCAATTCGATATTTC 5081
QY 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5082 AAAGACATGATGAGGCTTCAGAAAGACCATGACTGTTGAGGTAGAGGTTCCACTG 5139
RESULT 11
HSU81607 Homo sapiens 6608 bp mRNA linear PRI 26-JUN-1997
LOCUS Homo sapiens gravin mRNA, complete cds.
DEFINITION U81607
ACCESSION U81607.1 GI:2218076
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
```

source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
192..5537
/notes="kinase scaffold protein"
```

CDS

```
/codon_start=1
/product="gravin"
/protein_id="AAC51366.1"
/db_xref="GI:2218077"
/translation="MGAGSSTEQRPSEPEQSPGSGSPAEPEPSGGSPSAEAAPDPTADP
ATAASDPATKLLQKNGQLSTINGVAEQDELISLQEGDLNGKQKALNGALNGQALNGQSEBEV
IYTEVGQRDSERDSERDSKKNATKSAVVVDHLDIDGOENRNIEIIPSSSENLIELTQ
LQAGAAASKSEPTQSTKPEETLKEQSHAEISPPAESQGADEKECEKEQKEP
SKSASPTSTGTFKFTQCGWAGRKTKPRKPKGDEVAEAKSEKKQEPKPV
DPEBGKAEVASEKLVASQAHPAESHAEPRLSAEYKVELPSEEVSGSQGPE
EKPAFLATEVDEKLEHVEEVVAVHVTVEERTEOKTEVEETAGSVPAELVGM
APQSAEPAKELVKLKEVCVSGEDTQADLSPDEKLVSKPPEGVSVSEVMSQBRM
KVQGSPLKLLFTSLGKLSGKKQKGGGDESEHETQVPASFPKMTVPKRVRR
SSPEPEITCLEGLAEVQDGEAEAGTSDEKKREGVTPWASFKMVTTPKRVRR
PESDKDELDRKATLSSTESTASEMEQEMKGSVEPKPEPKRVKDVTSWDALEI
CVGSKKARRRSSDEGGPKAMGDHOKADEAGKDKETGDTGILAGSOEHDPGGS
SSPEOAGSPTGEGVSTWESFKELVTPRKKSLEKSEDSIAGSGVEHSTPDTEPG
KEESVSIKFFIPGRKKRPDGKQQAPEVDAGPTGANEEDSDVPAVFLSYDAVER
EKMEAQKQKGAEPQKAATEVSKELSESQVHMAAAVADGTRAATIIIEERSPSWIS
ASVTFPLEQVEAAALTEVLEREVIIEEPTVTEPLPENREARGDTPVSEAEITP
EAVTAAETAGPLGSEGTASAABETEMVASVOLTSDPTTEATPVQVEEGVPD
IEQERRTOEVLQVAEVKESQLPGTGDPEDVLPQVORAERPEEOAASGLAKE
TDVLVDAQEAKEPTFPQGVGOTTPESPEKAPQVTEIESSELVTTQAEETLAGV
KSQEMVQAAPPDSVEITPDSSETGTFVADFAGTTQKDEIWEIHENEVASGTQ
SGGTAEAVPAQKRPAPSPFQETKEQSKMBDTLEHTDKVESVETVSIKSTEG
TOEADQVADPEKTDVFPFEGLESIDGTITVSRKTEVALKEGTEEAECCKDDALE
LQSHAKSPSPVEREMVQVREKTEAETHVNEKLEHETAVTVSEEVSKOLLQTVN
VP1IDGAKEVSLGSPPCQLQGEAACTKIOVQSSSEASFTLTAAREEKVLGETANI
LETGTLSPAGAHVLEKSEKNEDEFAHPGEDAVTPGDCQAKSTPVIENATTKG
LSSDLEGTLSLKKSEDEVQACQVKSVALEDPENGILIELETSKSLVONI
IQTAVDVFRTEETATEMLTSELOAHVIKADSDAQOGETEKEBEPQASQDEPTPI
TSAKESESTAGVQAHSDIKSDMSAEKTMTEVEGTVNQDLVEEVLPSEEEGGG
AGTKSPVEDDGHALLAERIISKLVPKEDKGDVDDPENQNSALADTASGLTKES
PDTNGPKKEKEDAOEVLQEGKHSESDKAITTPQAQBELQKESAKSELTES"
```

BASE COUNT 2091 a 1418 c 1827 g 1272 t
ORIGIN

Alignment Scores:

Pred. No.:	1-85e-145	Length:	6608
Score:	4300.50	Matches:	967
Percent Similarity:	68.81%	Conservative:	189
Best Local Similarity:	57.56%	Mismatches:	408
Query Match:	53.27%	Indels:	116
DB:	9	Gaps:	36

US-09-902-432-4 (1-1596) x HSU81607 (1-6608)

QY	1	MetGlyAlaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp	19
Db	192	ATGGGCGCGGGAGCTCCACCGAGCAGCGAGCGCGCGCGCGAGGAGTCC	251
QY	20	ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla	38
Db	252	ACCGCGGCTGAGCCGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGAGCACC	311
QY	39	AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly	56
Db	312	ACCGCGGACCCCGCCATCGTGCCTCGGACCCCGCCACCGCCCTCTACAGAGATGGT	371
QY	57	GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsn	76
Db	372	CAGCTGTCCACCATCAATGGCTAGCTGAGCAAGATGAGCTCAGCTCCAGGAGGTGAC	431
QY	77	GlnGluGlyGln-----	82
Db	432	CTAAATGGCCAGAAAGAGGAGCCCTGAACGCTCAAGGAGGCCCTAAACAGCCAGGAGGA	491

QY 83 GIValValAspGluAspValGlyArgGluSerGluAspValArgGluLysAspArg 102
 Db |||||.....
 Db 492 GAAGTCATTGTCAAGGAGGTGGACAGAGACCTTGAAGAATGTGACCAAGAAGCTCC 551
 QY 103 ValGluMetAlaIleAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu 122
 Db |||||.....
 Db 552 GATTAAGATGGCTACTAAGTCAAGCGGTGTTCACGACATCAAGATGATGGGACGAG 611
 QY 123 GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluMetVal 142
 Db |||||.....
 Db 612 GAGAACCGGAAT--ATCAACACGATTCCTCTTCAGAAACGCAATTTAAGAGCTACA 668
 QY 143 GlnProAlaGluSerGlnIleAsnAspValGlyPheLysValPheLysPheValGly 162
 Db |||||.....
 Db 669 CAACCACTGAGTCCAGGCTAATGATTTGATTTAAGAGGTGTTTAAGTTTGTGGC 728
 QY 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGluLeuThr 182
 Db |||||.....
 Db 729 TTTAATTCACTGTGAAAAGATTAAGACAGAGAACCTGACCTGCCAGCTACTCACT 788
 QY 183 ValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaLysPheGlnGlu 202
 Db |||||.....
 Db 789 GTGAAGAAAGATCAAGGAGGAGCA-----GCAAGGGCTGTGCAACCAAGAGAG 839
 QY 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
 Db |||||.....
 Db 840 CCCAGCTT-----GGGCTGGAGAGAGACATCCAAAGAAAGCCAAACCATCT 893
 QY 223 ThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIleProLeuGln 242
 Db |||||.....
 Db 894 ACAGAAACCCAGAAAGACCTGAAAGCGTGAAGCAAGCAAGCAACCAATTTCTCCCCA 953
 QY 243 AlaGluSerAspGlnIleAlaGluGluGluAlaLysAspGluGlyGluGlnGlu 262
 Db |||||.....
 Db 954 GCCGATCTGGCCAAACAGAGT--GAGGAATGCCAAAGAGAGAGAGAAAGAAACAAGAA 1010
 QY 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
 Db |||||.....
 Db 1011 AAGAACTTACGAAAGTCTGCAATCTCCGACTGATCCGTGACCGATGAACAGGATCA 1070
 QY 283 SerPheLysLysPhePheThrHisGlyTyrAlaGlyTyrArgLysLysThrSerPheLys 302
 Db |||||.....
 Db 1071 ACCTTAAATAATTTCTTCACTCAAGGTGGCGCGCTGGCGCAAAAAGACATTTCAAG 1130
 QY 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
 Db |||||.....
 Db 1131 AAGCCAGAGAGATGAGTGAAGTCTTCAGAGAAAGAAAGAAAGAAAGCAAGCAAGAAAA 1190
 QY 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
 Db |||||.....
 Db 1191 GTTAGACACAGAAAGACGGAAGGACAGAGTTCCTCCGAGAACTGACCGCTCCGAG 1250
 QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
 Db |||||.....
 Db 1251 CAAGCCACCCACAGAGACCGGACAGAAAGTCCACAGAGCCCGCTTATAGTGAATAT 1310
 QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
 Db |||||.....
 Db 1311 GAGAAAGTTGAGTCCCTCAGAGAGCAAGTCAAGTGTCCGACGAGAACCTTCTGAAGAG 1370
 QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
 Db |||||.....
 Db 1371 AAACCTGCTCCGTGGCGACAGAAAGTGTGAGAGAAATAGAAAGTCCCAAGAAAGAG 1430
 QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGly 412
 Db |||||.....
 Db 1431 GTTGTGGCCGAATCCAGCTCAGACCGTGGAGAGAGAAACCAAGAGCG----- 1481
 QY 413 GlyLysAlaGluGlyGlyValValGluGlyThrGlyGluSerLeuProGluLys 432
 Db |||||.....
 Db 1482 -----AAACGGAAGTGAAGAAACAGACGGTCTGTGCCAGGTGAAGAA 1526
 QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452

Db 1527 TTGGTGAATGATGCAAGACCTCAGAGACCGAAGCTGCAAGAGAGCTGTAGACTC 1586
 QY 453 ArgGluMetCysValSerGlyLysPheThrGluLeuThrAspLeuSerProGluGlu 472
 Db |||||.....
 Db 1587 AAAGAAAGTGTGTTCGAGAGAGACCTTACAGAGAGGTGACCTTCACTCATATAG 1646
 QY 473 LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerGln 492
 Db |||||.....
 Db 1647 AAGGTGCTGCCAAACCCCGAAGGCTGTGAGTGAAGTGAAGTAATCTTCACTCAG 1706
 QY 493 GluArgLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
 Db |||||.....
 Db 1707 GAGAGAAAGAAAGTCCAGAGAAAGTCACTAAAGAAAGCTTTTACAGACACTGGCTTAAA 1766
 QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGly 532
 Db |||||.....
 Db 1767 AAGCTTCTCGAAAGAAACAGAAAGGAAAGA--GAGAGAGAGACGAGAAATCAGGG 1823
 QY 533 GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
 Db |||||.....
 Db 1824 GAGCACTCAGGTTCCAGCCGATTCGGAACAGCCAGAGAGCAAAAGAGGCGAGAGC 1883
 QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
 Db |||||.....
 Db 1884 TCTGCTCATCCCTGAGAGACCCAGAGAGATCACTGTCTCGAAAGAGCTTACCGCAG 1943
 QY 573 AlaProGlnAspGlyAlaGluGluGlyThrThrSerAspGlyLysLysArgGlu 592
 Db |||||.....
 Db 1944 GTGCAGCAGAGATGGGAGAGCTGAAGAAAGAGACTACTTCCGATGGAGAAAGAAAGAA 2003
 QY 593 GlyIleThrProThrAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
 Db |||||.....
 Db 2004 GGTGTACTCCTGGGCACTCATTTCAAAGATGTGAAGCCCAAGACGGTGTAGAGCG 2063
 QY 613 ProSerGluSerAspLysGluGluGluLysValLysSerAlaThrLeuSerSer 632
 Db |||||.....
 Db 2064 CTTTGGAAAGATTAAGAAAGATGACTGACAGAAAGTCAAGAGGCTTACTTGTCTTCC 2123
 QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluLys 652
 Db |||||.....
 Db 2124 ACCGAGAGACAGCTCTGAAATGCAGAAAGAAAGAAAGAGCGGTGAAGAGCCAAAG 2183
 QY 653 ProGluGluProLysArgArgValAspThrSerValSerTyrGluAlaLeuIleCysVal 672
 Db |||||.....
 Db 2184 CCGAAGAAACCAAGCGAAGGTGATCTCAGATCTTGGGAAAGCTTTAATTTGTGTG 2243
 QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArg 692
 Db |||||.....
 Db 2244 GGATCATCCAGAAAGAAAGCAAGAGAGAGTCTTCTGATGAGAGAGGGGACCAAAA 2303
 QY 693 ThrLeuGlyLysAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
 Db |||||.....
 Db 2304 GCAATGGAGAGAGACCAACAGAAAGCTGATGAGCGCGGAGAAAGAAAGAGAGCGGACA 2363
 QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGlu 732
 Db |||||.....
 Db 2364 GACGGGATCTCTTGTGTTCCCAAGAACATGATCCAGGCGAGGAGAAATCTTCCCGGAG 2423
 QY 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTyrGluSerPheLysArgLeu 752
 Db |||||.....
 Db 2424 CAAGCTGAAGCCCTTACCGAAGGAGGAGGCTTCTCAGCTGGAGATCTTTAAAGGTTA 2483
 QY 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysValGluAsp----- 769
 Db |||||.....
 Db 2484 GTACAGCCCAAGAAATAATCAAGTCCAAGTGAAGAGAAAGGAGAGACTCCATAGCT 2543
 QY 770 ---SerSerValGluGlnLeuSerThrGluLeuProSerArgGluGluSerTyrVal 788
 Db |||||.....
 Db 2544 GGGTCTGTGTGAACATTTCCACTCCAGACACTGAACCCGCTAAAGAAAGAAATCTGGGGTC 2603
 QY 789 SerIleLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808

Db	2604	TCAATCAAGAAAGTTTATTCTCGCAGAAAGGAAAGAAAGGCCAGATGGGAAAACAAGAACAA	2666
Qy	809	AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla	828
Db	2664	GCCTCTGTTGAGACGCGAGGCCCAACAGGGCCACAGAAGATGACTCTGATGTCTCCGGCC	2723
Qy	829	ValValProIeuSerGluThrAsnAlaValGluArgGluLysMetGlu-----Ala	845
Db	2724	GTGGTCCCTCTGTCTGAGTATGATGCTGTAGAAAGGAGAGAAATGGAGCCACAGCAAGCC	2783
Qy	846	GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer	865
Db	2784	CAAAAGCGCAGACAGCCCGGAGAGAGAGCGACCTCAGAGTGTCTCAAGAGACTCAGC	2843
Qy	866	LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer	885
Db	2844	GAGAGTCAGGTTCATATGATGGCAGCAGCTGTCTGTCAGCGACGAGCGAGCTACCAT	2903
Qy	886	ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr	905
Db	2904	ATTGAAGAAAGGCTCTCTTCTTGATATCTGCTCAGTGACAGAACCTCTTTGAACAAGTA	2963
Qy	906	AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu	924
Db	2964	GAAGCTCAAGCCGCACCTGTTAACTGAGGAGGATTGGAAAGAGAAAGTAA'TTTCGAAAGAA	3023
Qy	925	GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal	944
Db	3024	GAACCCCCACGGTTACTGAACCTCTGCAGAGAACAGAGAGCGCCGGGGCGACACCGCTC	3083
Qy	945	ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu	964
Db	3084	GTTAGTGAGGCGGAATTGACCCCGGAAGCTGTGACAGCTGCAGAAACTGCAGGGCCATTG	3143
Qy	965	ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla	984
Db	3144	GGTTCGAAGAGAAACCGAAGCATCTGCTGCTGAGAGAGACCACAGAAATGGTGTCAACA	3203
Qy	985	ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal	1004
Db	3204	GTCTCCAGTTNACCGACTCCACAGACCAACAGAGAGGCCACTCCGGTGCAGAGGTG	3263
Qy	1005	GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla	1024
Db	3264	GAAGTGGCGTACCTGACATAGAAGAGCAAGAGAGCGGACTCAAGAGGTCTCTCCAGCA	3323
Qy	1025	ValAlaAspLysValLysGluGluSerGlnValProAlaThr-----	1038
Db	3324	GTGGCAGAAAAGTGAAGAGGAATCCCACTGCTGCGCAGCGGTGGGCCAGAGAGATGTG	3383
Qy	1039	--GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu	1057
Db	3384	CTTCAGCCTGTGCAGAGACA-----GAGCGAGNAAGACCAAGAGAG	3425
Qy	1058	AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln	1077
Db	3426	CAGCTGAAGCGTCGGGTCTGAGAAGAGAGACCGATGTAGTGTTCGAAAGTAGAGTCTCAG	3485
Qy	1078	GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer	1097
Db	3486	GAGGCAAAACCTGAGCCCTTTTACACAAAGGAAGGTGTGGGGCGACACCACCCCAAGAAAGC	3545
Qy	1098	LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys	1113
Db	3546	TTTGAAAAGAGCTCTCAAGTCAACAGAGACGATAGAGTCCAGTGAGCTTTGTAACCACTTGT	3605
Qy	1114	GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla	1127
Db	3606	CAAGCCGAACCTTAGCTGGGGTAAATACAGGAGATGGTCAATGGAACAGGCTATCCCC	3665
Qy	1128	ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp	1147
Db	3666	CCTGACTCGGTGGAACCCCTACAGACGTGAGACTGTATGGAAGCACCCCGCTGAGCGAC	3725

Qy	1148	SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla	1167
Db	3726	TTTCGCACCGAGCCAAACCAGAAAAGCAGAGATTGTGGAAAAATCCATGAGGAGAATCGAG	3785
Qy	1168	ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu	1187
Db	3786	GTCGCATCTCGTACCAGTCAGGGGCCACAGAAAGCAGAGGCGAGTTCCTGCGCACAGAAAGAG	3845
Qy	1188	GluProSerThrLeuProAsnAenValProAlaGlnGluHisGlyGluGluProGly	1207
Db	3846	AGGCCTCCAGCA---CCTTCCAGTTTGTTGTTCCAGGAAGAACTAAAGACAATCAAAG	3902
Qy	1208	--ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu	1226
Db	3903	ATGGAAGACACTCTAGAGCATACAGATAAAGAGGTGTCAAGTGAACCTGATTCATTCTG	3962
Qy	1227	AlalysThrGluValGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys	1246
Db	3963	TCAAAGACTCAGGGAGCTCAA-----GAGCGTCGACCAATGTCGTGATGAGAAAAACCAA	4016
Qy	1247	-----GluGluGlnGluValPheValHisSerGly-----ProAsnSer	1259
Db	4017	GACGTACATTTTTCGAAGCACTTGAGGGCTCTATAGACACAGGCATTAACAGTCAGTCGG	4076
Qy	1260	GlnLysAlaAlaAspValThrTyAspSerGluValMetGlyValAlaGlyCysGlnGlu	1279
Db	4077	GAAAGGTCACTGAAGTTGCCCTTAAAGGTGAAGGCACAGAGAGCTGAATGTAAAAAG	4136
Qy	1280	LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu	1294
Db	4137	GATGATGCTCTGAACCTGCAGAGTCACGCTAAGTCTCTCCATCCCGCTGCAGAGAGAG	4196
Qy	1295	MetGluThrAspValGlnLysGlnLysArgGluThrLysProGluGlnValSerGluGlu	1314
Db	4197	ATGTTAGTTCAGTTCGAAAGGGGAGAAAACAGAAAGCAGAGCAACCCATGTGAATGAAGAG	4256
Qy	1315	Gly----GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrgLysProValLeu	1333
Db	4257	AAGCTTGAGCACGAAACAGCTGTACCGTATCTGAAGAGGTCAGTAAGCAGCTCCTCCAG	4316
Qy	1334	ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer---	1352
Db	4317	ACAGTGAATGTGCCCATCATAGATGGGCCAAAGGAAGTCAGCAGTTGTGAAGGAAGCCCT	4376
Qy	1353	ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp	1372
Db	4377	CCTCCCTCGCTAGGTCAAGAGGAGGCAGTATCCACCAAAATTCAGTTTCAGAGCTCTGAG	4436
Qy	1373	ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal	1390
Db	4437	GCATATTCTACTAACCGCGCTGCAGAGGAGGAAAAAGTGCTTAGGAGAAACTGCCAAC	4496
Qy	1391	IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys	1410
Db	4497	ATTTTAAAAACAGGTGAACGTTGGAGCCCTGCAGGTGCACATTTAGTTCTCGAAGAGAGAA	4556
Qy	1411	SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly	1430
Db	4557	TCCTCTGAAAAAATGAAGACTTTCGCGCTCATCCAGGGGAGAGATGCTGTGCCACAGGG	4616
Qy	1431	ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu	1450
Db	4617	CCCAGCTGTCCAGCAAAATCGACACAGTGATGATCTGCTATCCACAGAAAGGCTTGA	4676
Qy	1451	HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp	1470
Db	4677	AGTTCACACCTGGNAGGAGAGAAACCATCACTCACTGAAGTGAAGTCAGATGAAGTCAT	4736
Qy	1471	LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu	1490
Db	4737	GAGCAGGTGCTGTCGACAGAGGTC-----AAAGTGAGTGTAGCAATTCAGCATTTTA	4787

```

QY 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4788 -----GAGCCCGAAATGGAGATTGTGGAACTTGAGACCAAAAGCATTAACCTTGCCAA 4841
QY 1509 AsnValIleGluThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4842 AACATCATCCACACAGCCCTTGACAGCTTTGTACGTACAGAAAGAAACACCCAGCAAG 4901
QY 1528 HisAlaTyAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4902 TTGACGCTGAGATTACAGACACAGAGCTCAGCTGATTAAGCTGACAGCAGAGCGCTGA 4961
QY 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProAla 1562
Db 4962 CAGGAAACGAGAAAGAGAGAGAGAACTCAGGCTCTGCACAGGATGAAAACACCAAT 5021
QY 1562 LProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5022 ACTTCAGCCAAAGAGAGATCAGAGTCAACCGCAGTGGACAAAGCATTTCTATATTCC 5081
QY 1575 uGluuAtrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5082 AAAGACATGAGGAGGCTCAGAAAAGACCATGACTGTTGAGGTAGAGTTCACACTG 5139

RESULT 12
AR002603 AR002603 6605 bp DNA linear PAT 04-DEC-1998
LOCUS DEFINITION Sequence 4 from patent US 5741890.
AR002603 AR002603.1 GI:3964157
VERSION AR002603.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6605)
AUTHORS Scott J.D., Nauert J., Brian, and Kiauck, T.M.
TITLE Protein binding fragments of gravin
JOURNAL Patent: US 5741890-A 4 21-APR-1998;
FEATURES
source 1..6605
location/Qualifiers
BASE COUNT 2091 a 1416 c 1826 g 1272 t
ORIGIN

Alignment Scores:
Pred. No.: 9,21e-145 Length: 6605
Score: 4281.00 Matches: 965
Percent Similarity: 68.63% Conservative: 189
Best Local Similarity: 57.44% Mismatches: 409
Query Match: 53.03% Indels: 117
DB: Gaps: 37

US-09-902-432-4 (1-1596) x AR002603 (1-6605)
QY 1 MetGlyAlaGlySerSerThrGlnArgSerProGluGln---ProAlaGlySerAsp 19
Db 192 ATGGGGCGGGGAGTCTCCAGCGAGCCGACGCCGAGACGCCGCCAGAGGAGCTCC 251
QY 20 ThrProSerGluLeuValLeuSerGlnHisGlyProAlaAlaGluLaser---GlyAla 38
Db 252 AGCCCGGCTGAGCCGAGCCCGAGCGGGGGGCGCTCCGCGCAGAGCGGGCGCCAGAC 311
QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
Db 312 ACCGGGAGCCCGCCATCGCTGCTGCGACCCGCCACCAAGTCTTACAGAAATGGT 371
QY 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn 76
Db 372 CAGCTGTCCACCATCATGCTGAGCTGAGCAAGATGAGCTCAGCTCCAGAGGGGTGAC 431
QY 77 GlnGluGlyGln-----GlnGlu 82
Db 432 CTAAATGGCCAGAAAGAGCCCTTGAACGGTCAAGAGCCCTTAAACAGCCAGAGAGAA 491

```

```

QY 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
Db 492 GAAGTCATTTGTCACGAGAGTTTGACAGAGAGACTTGAAGTTGTGACCGAAAGAGACTCC 551
QY 103 ValGlnGluMetAlaAlaAsnSerThrAlaValAlaGluAspLysThrLysAspGlyGlnGlu 122
Db 552 GATAAAGATGGTACTACTAAGTCAAGCGGTTTCAACGATCAACATGATGATGGCGAGAG 611
QY 123 GlnThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetVal 142
Db 612 GAGAAACCGAAAT---ATCGAACAGATTCCTCTCAGAAACCAATTAAAGAGAGCTAAC 668
QY 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
Db 669 CAACCCACTGAGTCCCGAGCTTAATGATATTTGAAGAGGTGTTTAAGTTGTTGGC 728
QY 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuThr 182
Db 729 TTTAATTCCTGTGAAAGAGATTAAGACAGAGAGAGCTGACACTGTCCACTACTACT 788
QY 183 ValLysLysAspGluGlyGlyGlyAlaGluLaserValGlyAlaGlyAspHisGlnGlu 202
Db 789 GTGAAGAAAGATGAAGGGGAGAGCA-----GCAGGGCTGGCGACACACAGAGAC 839
QY 203 ProSerValGluThrAlaValAlaGlyGluSerLaserLysGluSerGluLeuLysGlnSer 222
Db 840 CCCAGCCCTT-----GGGGCTGGAGAACAGCATCCAAAGAAACGAAACCAATCT 893
QY 223 ThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerThrGlnIleProLeuGln 242
Db 894 ACAAGAAACCGAAGAGAGACCTTGAAGCTGACAAAGCCAGCAAGAAATTTCTCCCCA 953
QY 243 AlaGluSerAspGlnAlaAlaGluGlnGluLysAspGluGlyGlnGluLysGlnGlu 262
Db 954 GCCGATCTGGCCAAAGCAGTG---GAGGAATGCAAAAGAGAGAGAAAGAAACAAAGAA 1010
QY 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
Db 1011 AAAGAACTTACAGACTGACAGAAATCTCCGACTGATCCCGTGAACGAGAAACAGAGACA 1070
QY 283 SerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSerPheLys 302
Db 1071 ACCTTCAAAATAATTTCTCACTCAAGGTTGGCGGCTGGCCGCAAAACACAGATTTCAGG 1130
QY 303 LysSerLysGluAspAspLeuGlnThrAlaGluLysArgLysGlnGlnGluAlaGluLys 322
Db 1131 AAGCCGAAGAGAGATGAGTGAAGCTTCACAGAAAGAAAGAAAGAAAGACCAAGAAAA 1190
QY 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
Db 1191 GTAGACACAGAAAGAGAGAGAGAGAGAGTGGCTCCGAGAAACGTAGCCGCTCCAG 1250
QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTrp 353
Db 1251 CAAGCCACCCACAGAGACCGGAGAAAGTGCACAGAGCCCGGTTATACGCTGAATAT 1310
QY 354 GlnLysValGluLeuProLeuGlnValAlaGlyAspLeuGlnAlaSerSerGluGlu 373
Db 1311 GAGAAAGTTGAGCTCCCTCAGAGAGAGCAAGTCAAGTGCAGAGGACCTTCTGAAGAG 1370
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
Db 1371 AAACCTGCTCGGTGGCAGAGAAAGTGTGATGAGAAATAGAAAGTCCACCAAGAAAG 1430
QY 393 ValValAlaGluValHisValSerThrValGlnLysThrGlnGluGlnGlyGlyGly 412
Db 1431 GTTGAGCCGAAGTCAAGTCAAGCCGTGGAGGAGAGAAACGAAAGACAG----- 1481
QY 413 GlnGluAlaGluGlyGlyValValValGluGlyThrGlnLysLeuLeuProGluLys 432
Db 1482 -----AAAGAGAGGTGAGAAAGAAACAGCAGAGGCTGTGTCCAGCTGAAGAA 1526

```


QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
DB 1527 TTGTTTGAATGGATGATGAGAACCTTCAGAAAGCCGAACTGCGAAGGAGCTGGTGAAGCTC 1586
QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
DB 1587 AAAGAAACGTGTGTTTCCGAGAGGACCCATACAGAGGAGCTGACCTCAGTCTCTGATGAG 1646
QY 473 LysThrLeuProLysHisProGluGlyLeuValSerGluValGluMetLeuSerSerGln 492
DB 1647 AAGGTGCTGTCCAAACCCCGAAGCGCTGTGAGTGAGGTGGAAATGCTGTATCATCAG 1706
QY 493 GluArgIleValGlnGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
DB 1707 GAGAAATGAGGTGCAGGAGGTCCTAAAGAAAGCTTTTACCACTGCGCTTAAAA 1766
QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysGlyLysGluProGly 532
DB 1767 AAGCTTTCTGGAAGAAACAGAAAGGGAAGA---GGAGGAGGAGACGAGGAATCAGGG 1823
QY 533 GluThrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
DB 1824 GAGCACACTCAGGTTCACGCGGATTCCTCCGACAGCCAGGAGGAGCAAAAGGGCGAGAGC 1883
QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
DB 1884 TCTGCCCTCATCCCTGAGGAGCCGAGGAGATCATCGTGTCTGGAAGAGGCTTAGCCGAG 1943
QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
DB 1944 GTGCAGCAGATGGGGAAGCTGAAGAGGAGCTACTTCCGATGAGAGAAAAAAGAGAA 2003
QY 593 GlyIleThrProThrAlaSerPheLysLysMetValThrProLysLysArgValArg 612
DB 2004 GGTGTACTCCCTGGGCATCATTCAAAAGATGTGACGCCCAAGAAAGCGTTAGACGG 2063
QY 613 ProSerGluSerAspLysGluGluGluLysValLysLysValLysSerAlaThrLeuSerSer 632
DB 2064 CCTTCGAAAGTATAAAGAAAGATGAGCTGACAAAGTCAAGAGCGCTACCTGTGTCTTC 2123
QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
DB 2124 ACCGAGAGCAGCCTCTGAATGCAAGAAAGAAATGAAGGAGCGTGGAAAGCCAAAG 2183
QY 653 ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
DB 2184 CCGAAGAACCAAGCCAAAGTGTGATACCTCAGTATCTTGGGAAGCTTTAAATTTGTTG 2243
QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArg 692
DB 2244 GGATCATCCAAAGAAAGCAAGGAGGAGGAGGTCTCTTCTGATGAGGAGGGGAGCAAAA 2303
QY 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
DB 2304 GCAATGGGAGAGACCAACCAAGAAAGCTGATGAGCCGGAAGAAAGCAAGAGACGGGACA 2363
QY 713 AspAlaValProAlaSerThrGlnGlnAspGlnAlaGlnGlySerSerSerProGlu 732
DB 2364 GACGGATCTCTGCTGTTCCCAAGACATGATCCAGGGCAGGGAAGTTCCTCCCGGAG 2423
QY 733 ProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeu 752
DB 2424 CAAGCTGGAAGCCCTACCGAAGGGGAGGGGCTTTCCACCTGGGAGTCAATTTAAAGGTTA 2483
QY 753 ValThrProArgLysLysSerLysSerLysLeuGluLysAlaGluAsp----- 769
DB 2484 GTCACGCCAAGAAAAAATCAAGTCCAGCTCGAAGAGAAAAAGCGAAGACTCCATAGCT 2543
QY 770 ---SerSerValGlnGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpVal 788
DB 2544 GGGTCTGGTGTAGAACATTCCTCCTCAGACACTGAACCCCGTAAAGAAAGATCTCGGTGTC 2603
QY 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808

DB 2604 TCATCAAGAAAGTTTATCTCTGACGAAGGAAGAAAGGCCAGATGGGAAACAGAACAA 2663
QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
DB 2664 GCCCTGTTGAAGACGACGAGGCCCAACAGGGGCCCAACGAAGATGACTCTGTATGTCGCGGCC 2723
QY 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
DB 2724 GTGGTCCCTCTGTCTGATGATGATGCTAGAAAAGGGAGAAAAATGGAGGCACAGCAAGCC 2783
QY 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
DB 2784 CAAAAGGCCACAGACGAGCCCGAGCAGAGGAGCCACTGAGGTGTCCAAGGAGCTCAGC 2843
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
DB 2844 GACAGTCAGGTTCATATGATGGCAGCAGCTGTCTGCTGACGGAGGAGGCGAGCTACCAT 2903
QY 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
DB 2904 ATTGAAGAAAGGTCTCCTCTTGGATATCTGCTTCAGTGACAGAACCTCTTGAACAAAGTA 2963
QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
DB 2964 GAAGCTGAAGCCGACCTGTTAACTGAGGAGGTATTGGAAGAGAGAAATTTGCAGAAAGAA 3023
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
DB 3024 GAACCCCCACGGTTACTGAACCTCTGCCAGAGAACAGAGAGCCCGGGCGACACGGTC 3083
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
DB 3084 GTTAGTGAGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGCCCATTTG 3143
QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
DB 3144 GGTTCGGAAGAGAAACCGAAGCATCTGCTGCTGAAGAGACCACAGAAATGGTGTACGCA 3203
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluAlaThrProValGlnGluVal 1004
DB 3204 GTCTCCAGTTAACCCACTCCCGACAGACACAGAGAGGCCACTCGGTGTCAGGAGGTG 3263
QY 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
DB 3264 GAAGTGGCGTACTGACATAGAAGAGCAAGAGAGCGGAGCTCAAGAGGTCTCTCCAGGCA 3323
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
DB 3324 GTGCGCAAAAAGTGAAGAGGAATCCAGCTGCTGGCACCGGTGGGCCAGAGATGTG 3383
QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
DB 3384 CTTACGCTGTGCAGAGACA-----GAGCAGAAAGAACCCAGAAAG 3425
QY 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
DB 3426 CAGGCTGAAGCGTCGGTCTGAAGAAAGAGACGCGATGTAGTTGTTGAAGTAGATGCTCAG 3485
QY 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
DB 3486 GAGGCAAAAACCTGAGCCTTTTACACAAAGGGAAGGTGTGGGGCGAGACCACCCAGAAAGC 3545
QY 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
DB 3546 TTTGAAGAAAGCTCCTCAAGTCAAGAGAGCATAGATGCCAGTGTGTAACCATCTGT 3605
QY 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
DB 3606 CAAGCCGAAACCTTAGCTGGGTAAATACAGAGAGATGCTGATGGAAACAGGCTATCCCC 3665
QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147

```

Db 3666 CCTGACTGGTGAAGAAACCCCTTACAGACAGTGAAGTGAAGACCCCTTACCCAG 3725
Qy 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerIysAla 1167
Db 3726 TTGTGCGACGACGACCAACCCAGAAAGACGAGATTGTGGAATCCATGAGAGATGAG 3785
Qy 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnIysGlu 1187
Db 3786 GTGCATCTGTGA---CCAGTCAGGGGACAGAAAGACAGAGAGGCTTCTGCACAGAAAGG 3842
Qy 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGlnGluIleAspGluGluProGly 1207
Db 3843 AGGCTTCCAGCA---CCTTCCAGTTTGTGTTCCAGAGAAAGAACTAAAGACATCAAG 3899
Qy 1208 --ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProValLeu 1226
Db 3900 ATGMAAGACACTGAGATGACAGATAAGATAAGAGTGTCAGTGAACCTGATTCATTCTG 3959
Qy 1227 AlaIysThrGluValGlyGlnGluGluValAspThrPleuAspGlyGluIysValIys 1246
Db 3960 TCAAGAGCTGAGGGGAGCTCA---GAGGCTGACAGTATGCTGATGAGAAACCAAA 4013
Qy 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4014 GACGTACCATTTTTCGAAGAGACTTGAAGGCTGTATAGACACAGGCTAACAGTCACTGCG 4073
Qy 1260 GlnIysAlaAlaAspValThrTyraSerSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4074 GAAAGAGTCACTGAAGTTGCCCTTAAAGGTGAAGGACAGAAAGAGCTGAATGTAAAG 4133
Qy 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGluGlu 1294
Db 4134 GATGATGCTCTTGAACCTGAGAGTCAAGTAACTCTCCATCCGCCCTGAGAGAGAG 4193
Qy 1295 MetGluThrAspValGluIysGluIysArgGluThrIysProGluGlnValSerGluGlu 1314
Db 4194 ATGGTAGTTCAAGTGAAGAGGAGAAACAGAAACAGACCAACCATTTGATGAAGAG 4253
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluIleGluGluGluThrTyroIysProValLeu 1333
Db 4254 AAGCTGAGCAGAAACAGCTGTACCGTATCGAAAGAGTCAAGTACAGAGCTCTCCAG 4313
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyIysAlaLeuGlySerLeuGlyIysSer--- 1352
Db 4314 ACAGGAAATGTCCTCATCATGATGGGGCAAGAGTCAAGATTTGGAGAGAGCCCT 4373
Qy 1353 ProSerLeuProAspGlnAspIysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4374 CTTCCCTGCTAGTGTCAAGAGGAGGATGATGACCAAAATTCAGAGCTCTGAG 4433
Qy 1373 ThrThrValThrGlnThrAlaGluAlaValGluIysValIle-----GluThrValVal 1390
Db 4434 GCATCATTCACCTTCAACGCGGCTGACAGAGGAAAGAGTCTTGAAGAAACCTGCCAAC 4493
Qy 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluIys 1410
Db 4494 ATTTTAGAAACAGGTGAACCTTGGAGCTGCAAGTGCACATTTAGTTCTGGAGAGAA 4553
Qy 1411 SerSerAlaThrGlyGlyHisIleThrThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4554 TCCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGAAAGATGCTGTGCCACAGG 4613
Qy 1431 ProGluSerAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4614 CCCGACTGTCAGGCAAAATCCGACACCGATGATGATCTGCTACTCTCAAGAAAGCTTGA 4673
Qy 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
Db 4674 AGTTCGACCTGGAAGAGAGAAACCATCACTGAATGGAAAGTCAGATGAAGTCGAT 4733
Qy 1471 LysProAspAlaGlyProAspAlaAspGlyIysGluSerThrAlaIleGluIysValLeu 1490
Db 4734 GAGCAGGTTGCTTGCAGGAGGTC-----AAAGTAGTGATGCAATTAAGGATTTTA--- 4784

```

```

Qy 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerIysSerAsnIysIleValLeu 1508
Db 4785 -----GAGCCTGAAAAATGGGATTTTGGACCTTGAGACCAAAACAGATTAACCTGTCCA 4838
Qy 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4839 AACATCTCAGACAGACCGCTTGACAGTTTGTACGTACGAGAAAGAACCAACCGAAATG 4898
Qy 1528 HisAlaTyraSerSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4899 TTGACGTCTGAGTTACACACACAAAGCTCAGTGATTAAGCTGACAGCCAGAGCGCTGA 4958
Qy 1548 ArgCysIleThrIys-----MetLysAspAla-LysMetLysHisProVal 1562
Db 4959 CAGMAACGAGAAAGAGAGAGAGAACCTCAGGCTCTGCACAGATGAAGAAACCAAT 5018
Qy 1562 LProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5019 ACTTCAGCCAAAGAGAGACTCAGACTCAACCCAGTGGACAGACATTCGTGATTTTCC 5078
Qy 1575 uGluAlaThrAlaGlnProArgIysCysLeuProArgLeuGlnLeuIysAlaProVal 1594
Db 5079 AAAGCATGATGAGAAAGCTCAGAAAGAACCATGACTGTTGAGATGAAGGTTCCACTG 5136

RESULT 13
AF001504 6297 bp mRNA linear PRI 06-JUN-1997
LOCUS Homo sapiens myasthenia gravis autoantigen gravin mRNA, partial cds.
DEFINITION
ACCESSION AF001504
VERSION AF001504.1 GI:2160783
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 6297)
Bowlitch,R.D. and Ginsberg,M.H.
Sequence of gravin cDNA isolated from a human umbilical vein
endothelial cell library
JOURNAL
TITLE
AUTHORS
REFERENCE 2 (bases 1 to 6297)
Bowlitch,R.D. and Ginsberg,M.H.
Direct Submission
JOURNAL
TITLE
SUBMITTED (29-APR-1997) Biochemistry, University of Oklahoma Health
Sciences Center, PO Box 26901, BMSB 853, Oklahoma City, OK 73190,
USA

FEATURES
source
Location/Qualifiers
1..6297
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="endothelial cell"
/tissue_type="umbilical vein"
<1..5226
/note="kinase scaffold protein"
/codon_start=1
/product="myasthenia gravis autoantigen gravin"
/protein_id="AAB58938.1"
/db_xref="GI:2160784"
/translation="DPAIASDPATKLLQKNQGLSTINGABEDLSLQSGDLNGKQK
ALNQGALNQEEBEVIVTEVGQSDSDVERSDKMATKSAVAVHITDQGETPE
IIEIIPSESNLELTPTPSQANDIDIFKVFKEVGFVKDKDTEKPDVQLLVK
KDBEGAGAGDHDPSLGAEGASKESEPKOSTKEPEETLKREOSAEIISPASEGQ
AVBECKEGERKEKESKSAESPTSVTSSTGTFPKKFTFGQAGMRKRTSPKPKR
DEVASERKKKEQEKVDTBEDGAELVASEKLTSTSEBOAHQEPASHEMRPELSAEK
VELPSEVQSGSQSPSEKPAFLATEVDFDEKIEVHQBVAEVAHVSTVEKTEOKTE
VEETGAVPAEIVEMDAPEQAPAEVLVQLKTCVSGEDPTGADLSPDEKVLSTP
PEGVAVSEVMISQERMKVQSGPLKULFTSTGLKLSGKQKGRGGDBSGHTVQ
PADPSQOEQEGSSASPEPEITCLEKGLAEVQDGEAGATSDGKREGVAT
PMASFKMVTPKKRVRRPSPSDKDELIDYKSAITLSTSTESASMOEKGSVPEPKP
EERKRYDTISWALICVGS SKRARCGSSDDEGPKMGCHQKADKXCKETG
TDGILASQEHDPQSSSPQAGSPTEBGEVSTWESFKRLVTRKSKSKLBEKSED

```

STAGSGVEHSTPDTEPKESWWS IKKFIPIGRKKRPGDKQBPVEDAGPTGANNEDD
SDVPAVVPVLEDAVERKWEAQQAQSAEQPEQKAATEVSKELSESQVHMAAAVAD
GTAATIIIEERSPWTSASVTEPLEQVEAEALITBEVLEVEREVIAREEPTVTLELPE
NREAGDTTVSSEALTEPAVTAETAPGLGABEGTEASAAETTEMVSAVSQITDSDP
TTTEATPQVEVGGVPDIIEQERRTQEVQVAEKVBSQLPGTGGPBDVLQPVORA
EAERPEQAEASGLKKEATDVLKVDQAQEAETRPPTQGVGOTTTPESPEKAPQVTEI
ESSELVTCQAEITLAGVKSQWMEQAIIPDSVETPTDSETDGSFPVAFDARGTTQK
DRIVEIHENEVASGTQSGTEBAVPAQKRPAPSPSFQBEETKEQKMSMDLHET
KGEVSEVETVLIUSKTGTQADQYADEKTDVPFFEGLEGSIDTGITVSEKVEVAL
DKEGTBEACCKDQALQSHAKSPSPVERMVQVREKTEAEPTVHNEKLEBET
LTAAREEVKLGETANILETGETLEPAGAHVLEEKSSSEKNEDEFAHPGCEDAVPTGPD
LVTAAREEVKLGETANILETGETLEPAGAHVLEEKSSSEKNEDEFAHPGCEDAVPTGPD
COAKSTPVIATTKGLSDLEGEKTSILKWKSDREVQVACQEVKVSVAIEEDLEP
ENGILEETPSKLVQNIITQAVDQFVTEETATEMLTSELQTAHVIRKADSODAGQE
LTKGESETPQASQDEPTTSKAESESTAVGQAHSDISKDSEASEKTMVVEGSETV
NQDVEVLVPSSEEGGAGTKSPDEDDGHALLAERI BKSIVPEKDEKGDVDDPEN
ONSALADTDASGLTKESPDTPNGPKQEKEDAQVEVLQEGKVHSESDKAITPQAEEL
OKQERESAKSELTES"

BASE COUNT 2051 a 1310 c 1708 g 1228 t
ORIGIN

Alignment Scores:

Pred. No.: 5,17e-143 Length: 6297
Score: 4231.50 Matches: 943
Percent Similarity: 69.07% Conservative: 187
Best local Similarity: 57.64% Mismatches: 398
Query Match: 52.42% Indels: 108
DB: 9 Gaps: 31

US-09-902-432-4 (1-1596) x AF001504 (1-6297)

QY 41 AspProAlaThrLysLeuProGlnLysAenGlyGlnLeu 58
Db 1 GACCCCGCCATCGCTGCGACCCGCCCAACGCTCTCAGAGAATGTGTGACGTG 60
QY 59 SerSerValAenGlyValAlaGluGlnGlyAspValHisValGlnGluGluAenGlnGlu 78
Db 61 TCACCATCAATGCGGTAGCTGAGCAAGATGAGCTCAGCCTCCAGGAGGTGACCTAAAT 120
QY 79 GlyGln-----GluGluGluVal 84
Db 121 GGCCAGAAAGAGCCCTGAACGGTCAAGGACCCCTAAACAGCCAGGAGGAAGAAGTC 180
QY 85 ValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArgValGlu 104
Db 181 ATTGTTCACAGAGGTTCACACAGAGACTCTCAAGATGTGAGCGAAAGAGACTCCGATAAA 240
QY 105 GluMetAlaAlaAenSerThrAlaValGluAspIleThrLysAspGlyGlnGluGluThr 124
Db 241 GAGATGCTACTAAGTCAGCGGTGTTCACGACATCACATGATGGCGAGGAGGACAC 300
QY 125 SerGluIleIleGluGlnIleProAlaSerGluAenValGluGluMetValGlnPro 144
Db 301 CCCGAAATATCGACAGATTCCTCTTCAGAAAGCAATTTAGAGAGCTTAACACACCC 360
QY 145 AlaGluSerGlnAlaAenAspValGlyPheLysLysValPheLysPheValGlyPheLys 164
Db 361 ACTGAGTCCCAGGCTAATGATATGGATTTTAAAGAGGTGTTTAAAGTTTGTGGCTTAAA 420
QY 165 PheThrValLysLysAspLysAenGluLysSerAspThrValGlnLeuLeuThrValLys 184
Db 421 TTCACGTGTGAAAAGGATAAGACAGAGAAGCCTGACACTCTCCAGCTACTCTGTTGAAG 480
QY 185 LysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSer 204
Db 481 AAAGATGAAGGGAGGAGCA-----GCAGGGCTCGGCACCAACAGGACCCGAGC 531
QY 205 ValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLysGlnSerThrGlu 224
Db 532 CTT-----GGGGCTGGAGAAGCAGCATCCAAAGAAAGCGAACCCCAACAACTCTACAGAG 585
QY 225 LysGlnGluGlyThrLeuLysGlnGluLysSerThrGluIleProLeuGlnAlaGlu 244

Db 586 AAACCCGAAGAGACCCCTGAAGCGTAGCAAAAGCCAGCAAAATTTCTCCCCAGCGGAA 645
QY 245 SerAspGlnAlaAlaGluGluAlaLysAspGluGlyGlnGluLysGlnGluLysGlu 264
Db 646 TCTGGCAAGCAGTG--GAGGAATGCAAAAGAGGAAGAGAGAGAAACAAGAAAAAGAA 702
QY 265 ProThrLysSerProGluSerProSerProValAsnSerGluThrThrSerSerPhe 284
Db 703 CCTAGCAAGTCTCAGAATCTCCGAGTAGTCCCGTAGCCAGCCAGTGAAACAGGATCAACCTTC 762
QY 285 LysLysPhePheThrHisGlyTTPAlaGlyTTPAlaGlyLysThrSerPheLysLysSer 304
Db 763 AAAAAATTTCTTCACTCAAGTTGGCGCGCTGGCGCAAAAGAACCACTTTTCAGAGACCG 822
QY 305 LysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLysValAsp 324
Db 823 AAGGAGGATGAAGTGGAGCTTTCAGAGAGAGAAAAAGAACCAAGAGCCAGAAAAAGTAGAC 882
QY 325 GluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
Db 883 ACAGAAAGAGACGGAAGGAGGAGGTGTGCTCCGAGAAACTGACCGCTCCCGAGCAAGCC 942
QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLys 355
Db 943 CACCCAGAGGAGCCGCGAGAAAGTGCACAGCCCGGTTTATCAGCTGAATATGAGAAA 1002
QY 356 ValGluLeuProLeuGluAspGlnValGlyAspLeuAlaSerSerGluGluLysCys 375
Db 1003 GTTGAGCTGCCCTCAGAGGAGCAAGTCAGTGGCTCGCAGGAGACCTTCTGAAGAGAAACCT 1062
QY 376 AlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---GluValVal 394
Db 1063 GCTCCCGTTGGCGACAGAAGTGTGTGATGAGAAATAGAAAGTCCACCAAGAGAGGTGTG 1122
QY 395 AlaGluValHisValSerThrValGluLysThrGluGluGlnGlnGlyGlyGlyGlu 414
Db 1123 GCCGAAGTCCAGCTCAGACCGGTGGAGAGAGAACCAAGAGAGCAG----- 1167
QY 415 AlaGluGlyValValValGluGlyThrGlyGluSerLeuProGluLysLeuAla 434
Db 1168 -----AAACCGGAGGTGGAGAAACAGCAGGCTCTGTGCCAGCTGAAGAATTGGTT 1218
QY 435 GluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSerArgGlu 454
Db 1219 GAAATGGATGCAGAACCTCAGGAAGCCGAACCTGCCAAGAGCTGTGTGAAGCTCAAGAA 1278
QY 455 MetCysValSerClyGlyAspHisThrGlnLeuThrAspLeuSerProGluGluLysThr 474
Db 1279 ACGTGTGTTTCCGAGAGGAGCCCTACAGGGAGCTGACCTCAGTCTGTATGAGAAGGTG 1338
QY 475 LeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGlnGluArg 494
Db 1339 CTGTCCAAACCCCGAAGCGTGTGTGAGTGAGTGGAAATGCTGTCTCATCAGAGGAGA 1398
QY 495 IleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeu 514
Db 1399 ATGAAGTGCAGGGAAGTCCACTAAAGAGCTTTTACCAGCACTGGCTTAAAAAGCTT 1458
QY 515 SerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluProGluGlyGluTyr 534
Db 1459 TCTGGAAGAGAAACAGAAAGGAGAAAGAA---GGAGGAGGAGACAGGAATCAGGGGAGCAC 1515
QY 535 GlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerSerAla 554
Db 1516 ACTCAGGTTCCAGCCGATTTCTCCGACAGCCAGGAGAGCGAAGGCGGAGAGCTCTGCC 1575
QY 555 SerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPro 574
Db 1576 TCATCCCTGAGGAGCCCGAGGAGATCAGCTGTCTGGAAGAGGCTTAGCCGAGGTGACG 1635
QY 575 GlnAspGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGluGlyIle 594
Db 1636 CAGGATGGGGAGCTGAAGAAGGAGCTACTTCCGATGGAGAGAAAAAGAGAAAGGTGTC 1695

QY 595 ThrProTAlaSerPheLysLysMetValThrProLysLysArgValArgArgProSer 614
 Db 1696 ACTCCTGGGCTTCAATTCAAAAAGATGGAGCCCAAGAACCGTGTAGAGCGGCTTCG 1755
 QY 615 GluSerArgLysGluGluLysValLysSerAlaThrLysSerThrAsp 634
 Db 1756 GAAAGTGAATAAGATAGCTGGCAAGGTCMAAGCGCTACCTTGCTTCCACCGAG 1815
 QY 635 SerThrValSerGluMetGlnAspGluValLysThrValGluGluLysProGlu 654
 Db 1816 AGCAAGCCTCTGAATATGCAGAGAAATGAAGGAGCGTGGAAGACCAAGCCGGA 1875
 QY 655 GluProLysArgArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySer 674
 Db 1876 GAACCAAGCCGACAGTGTATCCTCAGTATCTTGGAAAGCTTTAATTGTGTGGATTA 1935
 QY 675 SerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGluGluProArgThrLeu 694
 Db 1936 TCCAAGAAAAGCAAGAGAGGCTCCTCTTGATGAGAGAGAGGAGACCAAAAGCAATG 1995
 QY 695 GlyGlyAspSerHisArgLysGluAlaSerLysAspLysGluAlaGlyThrAspAla 714
 Db 1996 GGAGAGAGACCAACAGAAAGCTATAGAGCCGGAAGAAAGCAAGAGACGAGACGAG 2055
 QY 715 ValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAla 734
 Db 2056 ATCCTTGCTGTCTCCCAAGAACATGATCCAGGCGAGGGAAGTTCTCCCGGAGCAAGCT 2115
 QY 735 GlySerProSerGluGluGluGluValSerThrTrpGluSerPheLysArgLeuValThr 754
 Db 2116 GGAACCCCTAACGAAAGGAGGAGGCGTTTCACCTGGAGTCAATTAAAGGTATGACAG 2175
 QY 755 ProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp-----Ser 770
 Db 2176 CCAAGAAAAAATCAAAAGTCCAAAGCTGGAAGAAAGAAAGCAAGACCTCAGTGGGTCT 2235
 QY 771 SerValGlnLysLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIle 790
 Db 2236 GGTGTAGAATTCACATTCACAGACCTGAAACCGGTAAAGAAATCCGTGGCTCAATC 2295
 QY 791 LysLysPheIleProGluArgArgLysArgAlaAspGluLysGlnGluGlnAlaThr 810
 Db 2296 AAGAAGTTATTCCTGGAGCAAGAAAGGCGCAGATGGAAACAAAGAACAGCCCT 2355
 QY 811 ValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAlaValAla 830
 Db 2356 GTTGAAGACGACAGGCGCAAGGCGCCACGAAGATGACTGTCCCGCGCTGATC 2415
 QY 831 ProLeuSerGluTrpAsnAlaValGluArgGluLysMetGlu-----AlaGlnGly 847
 Db 2416 CCTGTCTGATATGATCTGTAGAAAGGAGAAATGAGAGCACAGCAAGCGCAAAA 2475
 QY 848 AsnThrGluLeuProGlnLeuLeuGluAlaValTrpValSerGluGluLeuSerLysThr 867
 Db 2476 ACGCGAGACGCCCGCAAGAAAGCAAGCCACTGAGGTGTCCAAAGAGCTCAGCGAGAG 2535
 QY 868 LeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGlu 887
 Db 2536 CAGGTTCATATGATGACAGAGCTGTGCTGACGAGGACGAGGCACTACATATATGA 2595
 QY 888 GluArgSerProSerTrpLysSerAlaSerValThrGluProLeuGluLysIleThrAlaGly 907
 Db 2596 GAAAGTCTCTTCTTGATATCTGCTTCAGTACAGAACTCTTGAAACAAGTAGAGCT 2655
 QY 908 GluAlaMetProProValGluGluValThrGluLysAspIleIleIleIleIleIleIle 926
 Db 2656 GAAGCGCACTGTAACTAGAGGATATTGAAAGAGAAATTCAGAGAAAGAAAGCC 2715
 QY 927 ProValLeuThrGlnThrLeuProGluGluLysAspAlaHisAspMetValThrSer 946
 Db 2716 CCACGGTATGAACTCTGCGCAAGAACAGAGAGCCCGGAGCAACAGGTGCTTATG 2775

QY 947 GluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThr 966
 Db 2776 GAGCGGAATTGACCCCGAAGCTGTGACAGTGCAGAAATCGAAGGCCATGTGGTGC 2835
 QY 967 GluGluValThrGluAlaSerGlyValaGluGluThrThrAspMetValSerAlaValSer 986
 Db 2836 GAGAGGAGAACCGAAGCATCTGCTGTGAAGACACAGAAATGTGTCTCAGCACTTCC 2895
 QY 987 GlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValGluSer 1006
 Db 2896 CAGTTAACCGACTCCCGACAGACCAAGAGAGGCACTCCGCTGACAGAGATGGAAAGT 2955
 QY 1007 GlyValIleAspThrGlnGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAla 1026
 Db 2956 GAGTAACTGACATTAAGAGAGAGAGAGGAGGAGTCAAGAGTCTCCAGGCAAGTGCA 3015
 QY 1027 AspLysValLysGluGluSerGlnValProAlaThr-----Gln 1039
 Db 3016 GAAAAGTGAAAGAGATTCACAGAGCTGCTGGCACCGGTGGGCCAAGATGTGCTTCA 3075
 QY 1040 ThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluLysSer 1059
 Db 3076 CCTGTGCAGAGCA-----GAGCGACAAGACCAAGAACGACGAGCT 3117
 QY 1060 GluValIleLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAla 1079
 Db 3118 GAAGGTCGGCTTGAAAGAAAGACGAGATGATGTTGAAGTAGATCTCAGAGAGCA 3177
 QY 1080 GlyAlaGlnHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGlu 1099
 Db 3178 AAACTGAGCTTTTACACAAAGAGAGTGTGGGCGAGACCAACCCCAAGAACTTTGA 3237
 QY 1100 ---ValProGluValThrAlaAspValAspHis-----ValAlaThrCysGlnVal 1115
 Db 3238 AAAGCTCTCAAGTACAGAGAGCATAGATCCAGTGTATCACTTGTCAAGCC 3297
 QY 1116 -----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAlaProGlu 1129
 Db 3298 GAAACCTTAGCTGGGGTAAATCACAGAGAGTGTGAATGGAACAGGCTATCCCTTGA 3357
 QY 1130 SerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAsp 1149
 Db 3358 TCGGTGAAACCCCTACAGACAGTGAAGTGAAGAACACCCCGTATCCGACCTTTGAC 3417
 QY 1150 ThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAla 1169
 Db 3418 GCACAGACCAAAACCAAGAAAGACAGATTTGGAAATCCATGAGAGAAATGAGGTCCA 3477
 QY 1170 AlaValArgGlnSerGlnValThrGluGluGluAlaIleAlaIleAlaGlnLysGluLysPro 1189
 Db 3478 TGTGTACCCAGTCAAGGGGCGACAGAACAGAGGCACTTCTGCAACAGAAAGAGGCT 3537
 QY 1190 SerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly---Arg 1208
 Db 3538 CCAGCA-----CTTCCAGTTTGTGTTCCAGAGAAATCAAAAGAACATCAAAAGTGA 3594
 QY 1209 AspValLeuGluProThrGlnGlnGluLeuThrAlaIleAlaValProValLeuAlaLys 1228
 Db 3595 GACACTTAAAGCATTAACAGATTAAGAGGTGTCAAGTGAAGAACTGTATTCATCTGCAAG 3654
 QY 1229 ThrGluValGluGlnGluGluValAspTrpLeuAspGlyGluLysLys----- 1246
 Db 3655 ACTGAGGGGACTCAA-----GAGGCTACCAATGTGCTGATGAGAAACCAAAAGCGTA 3708
 QY 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSerGlnLys 1261
 Db 3709 CCATTTTGAAGAGACTTGAAGGGGTCTATAGACAGCATTAACGTACGTGGGAAGAG 3768
 QY 1262 AlaAlaAspValThrTrpAspSerGluValMetGlyValaGlyCysGlnGluLysGlu 1281
 Db 3769 GTCACTGAAGTTGCCCTTAAGAGTGAAGGACAGAAAGCTGAATGTAAAGAGTAT 3828
 QY 1282 SerThrGluValGlnSer-----LeuSerLeuGluGluGluLysMetGlu 1296

AL590413.18 GI:17426497
 HTG.
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 104939)
 Martin, S.
 Direct Submission
 Submitted (06-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Dec 8, 2001 this sequence version replaced gi:17384482.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; SW:
 SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep/ This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr6
 RP11-351K16 is from the library RPC1-11.2 constructed by the group
 of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-351K16. It may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true right end of clone RP11-351K16 is at 104939 in this
 sequence. The true left end of clone RP11-108N8 is at 59858 in this
 sequence. The true right end of clone RP3-403M6 is at 2000 in this
 sequence.

FEATURES	Location/Qualifiers
source	1..104939 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="6" /clone="RP11-351K16" /clone_lib="RPC1-11.2" 51026..51045 /notes="Sequence from AC023201 sequenced by WIBR."
misc_feature	51799..51952 /note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
misc_feature	52345..52447 /notes="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly." 86009..86146 /notes="Sequence from overlapping clone RP11-108N8 (AL590543). Assembly confirmed by restriction digest."
misc_feature	87770..87807 /note="Sequence from uni-directional dGTP big dye terminator reads only."

BASE COUNT 30065 a 22273 c 22882 g 29719 t
 ORIGIN
 Alignment Scores:

Pred. No.: 4.236-137 Length: 104939
 Score: 4097.00 Matches: 917
 Percent Similarity: 68.28% Conservative: 181
 Best Local Similarity: 57.03% Mismatches: 391
 Query Match: 50.75% Indels: 119
 DB: 9 Gaps: 32
 US-09-902-432-4 (1-1596) x AL590413 (1-104939)
 QY 77 GlnGluGluGlnGluGluGluValValAspGluAsp----- 88
 Db 9313 CAGCCGGGGCAAGAGTAGAGCCCTGTGGAGAAACAATTCTTGTATTGTAATCA 9372
 QY 89 -----ValGluGluArgGluSer 94
 Db 9373 CCTTTCTCTCTCCACCCGCCCTTTTGTATATAGTAGACAGAGACTCT 9432
 QY 95 GluAspValArgGluLysAspArgValGluGluMetAlaAlaAspSerThrAlaValGlu 114
 Db 9433 GAGAGTGTAGCAAAAGAGACTCCGATAAAGATGCTATCAAGTCACGGCTTGTAC 9492
 QY 115 AspIleThrLysAspGluGlnGluGluThrSerGluIleIleGluGlnIleProAlaSer 134
 Db 9493 GACATCAAGATATATGGCAGAGAGAGACACCCGAAATATTCGAACAGATTCCTTCA 9552
 QY 135 GluAsnAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPhe 154
 Db 9553 GAAAGCAATTTAGAAAGAGCTAACACAAACCACAGTCCAGCTCATATATATGATTT 9612
 QY 155 LysLysValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGluLys 174
 Db 9613 AAGAAAGTGTAAAGTTGTGGCTTAAATCTAGTGAAGAAAGATTAAGACAGAGAAAG 9672
 QY 175 SerAspThrValGlnLeuLeuThrValLysLysAspGluGluGluValAlaGluAlaSer 194
 Db 9673 CCTGACACTCTCAGCTACTCACTGTGAAGAAAGATGAAGGGAGGAGCA----- 9723
 QY 195 ValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGluGluSerAlaSer 214
 Db 9724 GCGAGGGCTGTGGCAGCAACCAAGACCCAGCCTT-----GGGGCTGAGAGAGCATCC 9777
 QY 215 LysGluSerGluLeuLysGlnSerThrGluLysGlnGluGluGluGluGluGluGln 234
 Db 9778 AAGAAAGGCAACCAACATCTACAGAAACCCGAAAGAGCCTGAAAGCTGTAGCAA 9837
 QY 235 SerSerThrGluLeuProLeuGlnAlaGluSerAspGlnAlaGluGluGluAlaLys 254
 Db 9838 AGCCACGCAAAATTTCTCCCGACCGAATCTGGCCAGCACTG---GAGGAATGCATA 9894
 QY 255 AspGluGluGluGluLysGlnGluLysGluProThrLysSerProGluSerProSerSer 274
 Db 9895 GAGCAAGGAGAAAGAAACAAAGAAAGAAAGAAAGCTGACAGATCTCCGACTAGT 9954
 QY 275 ProValAsnSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTTPAlaGly 294
 Db 9955 CCGGTACCAAGTAAACAGAGATCAACCTTCAAAAAATTTCTCACTCAAGGTTGGCCGGC 10014
 QY 295 TrpArgLysLysThrSerPheLysLysSerLysGluAspAspLeuThrAlaGluLys 314
 Db 10015 TGGCGCAAAAGACCAAGTTTCAGGAACCCGAAGAGATTAAGTGAAGCTTCAGAGAAAG 10074
 QY 315 ArgLysGluGlnGluAlaGluLysValAspGluGluGluLysGluLysThrGluProAla 334
 Db 10075 AAAAAGAAACAAAGCCAGAAAGTAAGACACAGAAAGAAAGCGAAGAGGTGGCC 10134
 QY 335 SerGluGlu-----GlnGluProAlaGluAspThrAsp 345
 Db 10135 TCCGAGAAACCTGACCGCTCCGAGCAAGCCACCAAGAGAGCCGGCAGAAAGTGCAC 10194
 QY 346 GlnAlaArgLeuSerAlaAspArgLysLysValGluLeuProLeuGlnGluAspGlnValGly 365
 Db 10195 GAGCCCGGTTATCAGCTGAATATGAGAAAGTTGAGCTGCCCTCAGAGAGCAAGTCACT 10254

QY 366 AspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGlu 385
 Db 10255 GAGCTCGAGGAACTTCTGAAGAGAAACCTCTCCGTGGGAGCAAGATTTGATGAG 10314
 QY 386 LysMetGluAlaHisGln---GluValValAlaGluValHisValSerThrValGluLys 404
 Db 10315 AAAATGAAAGTCCACCAAGAAAGAGGTTGTGCCGGAAGTCCAGTCAAGCCGTGAGAG 10374
 QY 405 ThrGluGluGluGlnGlyGlyGlyGlyGluAlaGluGluGlyValValGluGlyThr 424
 Db 10375 AGAACCGAAAGACAG-----AAACGAGAGTGAAGAAACA 10410
 QY 425 GluGluSerLeuProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGlu 444
 Db 10411 GCGAGTCTGTGCGCACTGAAGAAATGGTTGAATGATGACAGAACTTCAGAGAGCTTAA 10470
 QY 445 ProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGln 464
 Db 10471 CCTGCCAAGAGCTGTGTGAACCTCAAGAAACGTGTCTTCGAGAGAGACCTTACACAG 10530
 QY 465 LeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGluGlyIleValSer 484
 Db 10531 GAGCTGACCTCAGCTCATGATGAGAAAGTGTCTGTCAAACCCCGAAGAGCGTTGTAGT 10590
 QY 485 GluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLysLys 504
 Db 10591 GAGGTGAAATGCTGTATCAAGAGAGAAATGAAGGTGCAGAGGAAGTCACTAAAGAAAG 10650
 QY 505 LeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGly 524
 Db 10651 CTTTTCACAGACACGCTTAAAGAAAGCTTCTGGAAGAAACAGAAAGGAGAAAG--- 10707
 QY 525 GlyGlyGlyAspGluGluProGluGlyGluGlyGlnHisIleHisThrGluSerProGluSer 544
 Db 10708 GAGAGAGAGAGAGAGAAATCAAGAGAGACACTCAGTGTCCAGCCGAAATTCCTCGAGAC 10767
 QY 545 AlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluProGluGluThrThr 564
 Db 10768 CAGAGAGAGCAAAAGAGGAGAGCTGTGCTCATCTCCCTGAGAGACCCGAGAGATCAAG 10827
 QY 565 CysLeuGluLysGlyProLeuGlnAlaProGlnAspGlyGluAlaGluGluGlyThrThr 584
 Db 10828 TGTCTGAAAGAGGCTTTCAGAGTGCAGACAGATGGGAAAGCTGAAAGAGACTACT 10887
 QY 585 SerAspGlyGluLysArgGluGlyIleThrProTTPAlaSerPheLysLysMetVal 604
 Db 10888 TCCGATGAGAAAGAAAGAAAGAGATGATCACTCCCTGGGCACTCATTAAGAAAGATGGTG 10947
 QY 605 ThrProLysLysArgValArgArgProSerGluSerAspLysGluGluGluLeuGluLys 624
 Db 10948 ACGCCCAAGAACGTTTATGACAGGCTTCGAAAGTGAATGAAGATGAGCTGAGCAAG 11007
 QY 625 ValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluVal 644
 Db 11008 GTCAAGAGGCTTACTTGTCTTCCACAGAGACACAGCTCTGAAATGCAAGAAAGAAAG 11067
 QY 645 LysThrValGlyGluGlnGlnLysAspProGluGluProLysArgArgValAspThrSerVal 664
 Db 11068 AAAGGAGCGTGAAGAGCCAAAGCCGGAAGAAACCAAGCCGCAAGGTGATACCTCAGTA 11127
 QY 665 SerTPGluAlaLeuIleLysValGlySerSerLysArgAlaArgLysAlaSerSer 684
 Db 11128 TCTTGGAGGCTTTAATTTGTGTGATCATCAAGAAAGAGCAAGAGAGAGGTCCTT 11187
 QY 685 SerAspAspGluGlyGlyProArgThrLeuGlyLysSerHisArgAlaGluAla 704
 Db 11188 TCTGATAGAGAGAGGAGACCAAAAGCAATGGAGAGACCAACCAAGACTATAGAGCC 11247
 QY 705 SerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGln 724
 Db 11248 GGAAAGACAAAGAGACGAGGAGACAGCGGATCTTGTGCTTCCAAAGAAACATGATCA 11307
 QY 725 AlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGluGluGlyValSer 744

Db 11308 GGGCAGGGAAGTTCCTCCCGGAGCAGCTGGAAAGCCCTACCGAAGGGGAGGGCGTTTCC 11367
QY 745 ThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLysLysLeuGlu 764
Db 11368 ACCTGGAGTCAATTAAAGGTTAGTCACGCCAAGAAAAATCAAAAGTCCAAGCTGGAA 11427
QY 765 GluLysAlaGluAsp-----SerSerValGluGlnLeuSerThrGluIleGlu 780
Db 11428 GAGAAAGCGAAGACTCCATAGCTGGGTCTGGTGTAGAACATTCCACTCCAGACACTGAA 11487
QY 781 ProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArgLysLys 800
Db 11488 CCCGTTAAGAAGAACTCTGGGTCTCAATCAAGAGTTTATCTGGACCGAGAGAGAAA 11547
QY 801 ArgAlaAspGlyLysGlnGlnAlaThrValGluAspSerGlyProValGluIleAsn 820
Db 11548 AGCCACATGGGAAACAAGAACAGCCCTGTTGAAGACGCGAGGCCCAACAGGGGCCAAC 11607
QY 821 GluAspAspProAsnValProAlaValValProLeuSerGluTrpAsnAlaValGluArg 840
Db 11608 GAAGATGACTCTCTGCTCCCGCGGTGCTCCTCTGCTCAGTATGATGCTGTAGAAAGG 11667
QY 841 GluLysMetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAla 857
Db 11668 GAGAAATGGAGGCACAGCAGCCCAAAAGCGCAGAGCGCCCGGAGAGAGCGAGCC 11727
QY 858 ValTyrValSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIle 877
Db 11728 ACTGAGGTGTCCAAGGAGCTCAGCGAGAGTCAGGTTTCATATGATGGCAGCAGCTGTGCT 11787
QY 878 AspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSer 897
Db 11788 GAGCGGACGAGGGCAGCTACCATTAATTGAAGAAAGGTCTCCTCTTGGATATCTGCTTCA 11847
QY 898 ValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThr 917
Db 11848 GTGACAGAACTCTTTGAAACAGTAGAAGCTGAAGCCGCACCTGTTAACTGAGGAGGTATTG 11907
QY 918 GluLysAspIleIleAla----GluGluThrProValLeuThrGlnThrLeuProGluGly 936
Db 11908 GAAAGAGAAAGTAATTGACAGAAAGAAACCCGCCAGCTTACTGAACCTCTGCCAGAGAAC 11967
QY 937 LysAspAlaHisAspAspMetValThrSerGluValAspPheThrSerGluAlaValThr 956
Db 11968 AGAGAGCCCGGGCGACACGCTGTTAGTGAGCGGAAATTGACCCCGAAGCTGTGACA 12027
QY 957 AlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGlu 976
Db 12028 GCTGCAGAAACTCGAGGCCATTGGGTGCCGAGAGAAAGGAACCGAAGCATCTGCTGTGAA 12087
QY 977 GluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGlu 996
Db 12088 GAGACACACAGAAATGGTGTGACGAGTCTCCAGTTAAACCGACTCCCGACACACAGAG 12147
QY 997 GluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluArg 1016
Db 12148 GAGGCCACTCCGGTGCAGGAGGTGGAAGTGGGGTACTCTGACATAGAAGAGCAAGAGAGG 12207
QY 1017 GlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValPro 1036
Db 12208 CGGACTCAAGAGTCTCCAGGCGAGTGGCAGAAAGTGAAGAGGGAATCCCGAGCTGCCT 12267
QY 1037 AlaThr-----GlnThrValGlnArgThrGlySerLysAlaLeu 1049
Db 12268 GGCACCGGTGGCCAGAGATGTCTTACGCTGTGACAGAGACA----- 12312
QY 1050 GluLysValGluGluValGluGluAspSerGluValLeuAlaSerGluLysLysAsp 1069
Db 12313 ---GAGGCAGAAAGACAGAGAGCAGCGCTGAAGCGTCGGGTCTGAAGAAAGAGACGGAT 12369
QY 1070 ValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGlu 1089
|||||

Db 12370 GTAGTCTTGAAGTAGATGCTCAGGAGGCCAAAACCTGAGCCTTTTACACAAGGGAAGGTG 12429
QY 1090 ThrGlyGlnAlaThrProGluSerLeuGlu---ValProGluValThrAlaAspValAsp 1108
Db 12430 GTGGGCGACAGCACCCCGAGAAAGCTTTGAAAAGCTCTCTCAAGTCACAGAGCAGATAG 12489
QY 1109 His-----ValAlaThrCysGlnVal-----IleLysLeuGlnGln 1120
Db 12490 TCCAGTGAAGCTTGTGTAAACCACCTTGTCAAGCCGAAACCTTAGCTGGGTAAATCACAGGAG 12549
QY 1121 Leu---MetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThr 1139
Db 12550 ATGGTCATGGACAGCTATCCCCCTGACTCGGTGAAAACCCCTACAGACAGTGAAGT 12609
QY 1140 AsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThr 1159
Db 12610 GATGGAAGCACCCCTAGCCGACTTTCGACGACACGAGGACCAACCCAGAAAGACGAGATT 12669
QY 1160 IleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGlu 1179
Db 12670 GTGGAATCCATGAGGAGAAATGAGTCCGATCTGTGATCCAGTCAAGGGGCGCACAGAAGCA 12729
QY 1180 GluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGln 1199
Db 12730 GAGGCAGTTCCTGCACAGAAAGAGAGCCCTCCAGCA---CCTTCCAGTTTGTGTTCCAG 12786
QY 1200 GluGluHisGlyGluGluProGly---ArgAspValLeuGluProThrGlnGlnGluLeu 1218
Db 12787 GAAGAAACTAAAGAACAAATCAAAAGATGGAAGACACTCTAGAGCATACAGATAAAGAGGTG 12846
QY 1219 ThrAlaAlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAsp 1238
Db 12847 TCAGTGGAAACTGTATCCATTCTGTCAAAGACTGAGGGGACTCAA-----GAGGCTGAC 12900
QY 1239 TrpLeuAspGlyGluLysValLys-----GluGluGlnGluValPheVal 1253
Db 12901 CAGTATGCTGATGAGAAAACCAAGAGCTACCATTTTTCGAAGGACTTGAGGGGTCTATA 12960
QY 1254 HisSerGly-----ProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGluVal 1271
Db 12961 GACACAGGCATAACAGTCAGTCGGGAAAAGGTCACCTGAAGTTGCCCTTAAAGGTGAAGGG 13020
QY 1272 MetGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSer----- 1287
Db 13021 ACAGAAAGAGCTGAATGTAAAAAGGATGATGCTCTTGAATCGCAGAGTCACGCTAAGTCT 13080
QY 1288 ---LeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThr 1306
Db 13081 CCTCCATCCCGTGGAGAGAGATGTTAGTTCAAGTCCAAAGGGGAGAAACAGAGCA 13140
QY 1307 LysProGluGlnValSerGluGly---GluGlnGluThrAlaAlaProGluHisGlu 1325
Db 13141 GAGCCAAACCATGTGAATCAAGAGAGCTTGAGCACGAAACAGCTGTACCGTATCTGAA 13200
QY 1326 GlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAla 1345
Db 13201 GAGGTCAAGTAAGCAGCTCCTCCAGACAGTGAATGTGCCCATCATAGATGGGCGAAAGGAA 13260
QY 1346 LeuGlySerLeuGlyLysSer---ProSerLeuProAspGlnAspLysAlaGlyCysIle 1364
Db 13261 GTCAGAGTTTGAAGAAAGCCCTCCTCCCTGCTAGGTCAAGAGGAGGAGTATGCACC 13320
QY 1365 GluValGlnValGlnSerLeuAspThrValThrGlnThrAlaGluAlaValGluLys 1384
Db 13321 AAAATTCAAGTTCAGAGCTCTGAGGCATCATCTCACTTAAACAGCGCTCGAGAGGAGAA 13380
QY 1385 ValIle-----GluThrValValIleSerGluThrGlyGluSerProGluCysValGly 1402
Db 13381 AAGGTCTTAGGAGAACTGCCAACATTTTAGAAACAGGTGAACCGTTGAGCCCTGCAGGT 13440
QY 1403 AlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHis 1422
Db 13441 GCACATTTAGTTCTGGAAGAGAAATCCTCTGAAAAAAAATGAAGACTTTGCCCGCTCATCCA 13500

Db 87826 AGCCACGCAAAATTTCTCCCGACCGCATCTGGCCAAAGCATGT--GAGCAATGCATA 87882
QY 255 AaPgluGlulGlulysGlnGluProThrLysSerProGluSerProSerSer 274
Db 87883 GAGCAAGGAGAAAGCAAGAAAGAAAGAAAGCAAGCAAGCTGCGAATCTCCGATAGT 87942
QY 275 ProValAspSerGluThrThrSerSerPheLysPhePheThrHisGlyTPAlaGly 294
Db 87943 CCGGTACCGAGTAAACAGAGATCAACCTTCAAAAATTTCTTCACTCAAGGTTTGAGCCGCG 88002
QY 295 TPAlaGlyLysThrSerPheLysSerLysGluAspAspLeuGluThrAlaGluLys 314
Db 88003 TGGCCGAAAAAGACCGCTTTCAGGAAGCCGAGAGAGATGAAGTGAAGCTTCAGAGAG 88062
QY 315 ArgLysGluGlnGluAlaGluLysValAspGluGluGluLysGluLysThrGluProAla 334
Db 88063 AAAAAGGAACAAGAGCCAGAAAAAGTAGACACAGAAAGACCGAAAGGCGAGGTTGCG 88122
QY 335 SerGluGlu-----GlnGluProAlaGluAspThrAsp 345
Db 88123 TCCGAGAAAACTGACCGCTCCGACGCAAGCCACCCACAGAGCGCGAGAAAGTCCCCAC 88182
QY 346 GlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGly 365
Db 88183 GAGCCCCGCTTATCACTGATATGAGAAAGTTGAGCTGCTCCAGAGAGCAAGTCAGT 88242
QY 366 AspLeuGluAlaSerSerGluLysCysAlaProLeuAlaThrGluValPheAspGlu 385
Db 88243 GCCTCCAGAGCACTTCTGAGAGAAACCTGCTCCGTCGCGACAGAGGTTGATGAG 88302
QY 386 LysMetGluAlaHisGln---GluValValAlaGluValHisValSerThrValGluLys 404
Db 88303 AAAATGAAGTCCACCAAGAGAGGTTGTGGCCGAAGTCCAGCTCCAGCCGTGAGAG 88362
QY 405 ThrGluGluGlnGluGlnGlyGlyGlyGluAlaGluGlyValValValGluGlyThr 424
Db 88363 AGAAGCGAAGAGAG-----AAAAAGAGGTGGAAGAAACA 88398
QY 425 GlyLysLeuProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGlu 444
Db 88399 GCAGGCTGTGTCCAGCTCAAGAAATGTTGAATGATGCAGAACCTCCAGAACTGAA 88458
QY 445 ProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGln 464
Db 88459 CCGCAAGAGAGTGTGAAGCTCAAGAAAGAGCTGTGTTCCGAGAGACCTTACAACA 88518
QY 465 LeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGluGlyLeuAlaSer 484
Db 88519 GAGAGTGAAGCTCAAGTCTGATGAGAGAGGTCTGTCCAAACCCCGAAGCGTGTGAGT 88578
QY 485 GlnValGluMetLeuSerSerGlnGluArgLeuLysValGlnGlySerProLeuLysLys 504
Db 88579 GAGGTGAATGTCTCATCAACAGAGAGATAGAGTGCAGAGAACTCCACTAAAGAAAG 88638
QY 505 LeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnGlyLysAspGly 524
Db 88639 CTTTTCACGACACTGCTTAAAGAGCTTCTGAAAGAAACAGAAAGGGAAGAAG--- 88695
QY 525 GlyGlyGlyAspGluGluProGluGlyGlyLysGlnHisLeuHisThrGluSerProGluSer 544
Db 88696 GAGAGAGAGAGACAGAGATAGAGGAGAGACACACAGGTTCCAGCCGATTCGAGACAGC 88755
QY 545 AlaAspGluGlnLysGlyLysSerSerAlaSerSerProGluGluProGluGluThrThr 564
Db 88756 CAGGAGAGAGAAAGGCGAGAGCTGTGCTTCACTCCCTAGAGAGCCGAGAGAGATCAGC 88815
QY 565 CysLeuGluLysGlyProLeuGluAlaProGluAspGlyGluAlaGluGluGlyThrThr 584
Db 88816 TGTCTGAAAAAGGCTTACCGCAGGTGCAGACAGATGGGGAACTGAAAGAGAGCTACT 88875
QY 585 SerAspGlyGluLysLysArgGlyGlyLeuThrProThrAlaSerPheLysMetVal 604
Db 88876 TCCGATGAGAGAAAAAGAAAGGTGTCTCTCTGGGCACTTCAAAAAAGATGTGTG 88935

QY 605 ThrProLysArgValArgArgProSerGluSerAspLysGlnGluGluLys 624
Db 88936 AGCCCAAGAAAGCGGTGAGAGCGCTTCGAAAGATGTATTAAGAAAGAGTGTGAGCAAG 88995
QY 625 ValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluVal 644
Db 88996 GTCAAGACGCTACTTGTCTTCCACGAGACAGAGCTCTGAAATCCAGAAAGAAATG 89055
QY 645 LysThrValGlyGluGluGlnLysProGluGluProLysArgValAspThrSerVal 664
Db 89056 AAAAGGACGTGGAAGAGACCAAGCCGAGAGAACCAAGCCGAGAGTGAATCCTCAGTA 89115
QY 665 SerTPAlaLeuAlaLeuLysCysValGlySerSerLysArgAlaArgLysAlaSerSer 684
Db 89116 TCTTGAGAACTTAAATTTGTGTGGATCATCCAAAGAAAGAGCAAGAGGCTCTCT 89175
QY 685 SerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAla 704
Db 89176 TCTGATGAGGAAGGGGAGCAAAAGCAATGGAGAGAGACACCAAGAAAGCTGATGAGCC 89235
QY 705 SerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGln 724
Db 89236 GGAAGAAAGCAAGAGACGGGGACAGACGGGATCTTGTCTGTTCCCAAGAAACATGATCCA 89295
QY 725 AlaGlnLysSerSerSerProGluProAlaGlySerProSerGluGlyGluLysValSer 744
Db 89296 GGGCAGGAGAGTCTTCCCGAGACAGCTGGAACCTTACCGAAGGGAGAGGCTTTTCC 89355
QY 745 ThrTPGluSerPheLysArgLeuValThrProGlyLysLysSerLysSerLysLeuGlu 764
Db 89356 ACCTGGAGTCACTTTAAAGTTACTACGCGCAAGAAAAATCAATCCAAAGCTGGA 89415
QY 765 GlyLysAlaGluAsp-----SerSerValGluGlnLeuSerThrGluLeu 780
Db 89416 GAGAAAGCGAAGACTCCATGACTGGTCTGGTGTAGAAACATTCACCTCCAGACACTGA 89475
QY 781 ProSerArgGluGluSerThrTPValSerLeuLysPheLeuProGlyArgArgLysLys 800
Db 89476 CCGGTAAAGAAATCTCGGTCTCAATCAAGAAATTTATCTCGAGAGAAAGAA 89535
QY 801 ArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyProValGluLeuAsn 820
Db 89536 AGGCCAGATGGGAAACAGAAACAGACCCCTGTGAAGACGAGGCCCAACAGGGGCCAAC 89595
QY 821 GluAspAspProAsnValProAlaValValProLeuSerGlyTyrAsnAlaValGluArg 840
Db 89596 GAAAGTGACTGTGATGTCCCGCGCTGTCTGTCTGATGATGATGTGTAGAAAGG 89655
QY 841 GlyLysMetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeuGluAla 857
Db 89656 GAGAAATGAGAGCAACAGCAAGCCAAAGAGCGAGAGAGAGCCGAGAGAGAGCGAGC 89715
QY 858 ValTyrValSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIle 877
Db 89716 ACTGAGTGTCCAAAGAGTCAAGAGTCAAGAGTCAATATGATGAGAGCAAGCTGTGCT 89775
QY 878 AspGlyThrArgAlaValThrSerValGluGluArgSerProSerThrPheSerAlaSer 897
Db 89776 GAGGGAGAGAGGAGCTTCACTTATGGAAGAAAGTCTCTTGTGATATCTCTCTCA 89835
QY 898 ValThrGluProLeuGlnHisThrAlaGlyAlaMetProProValGluGluValThr 917
Db 89836 GTGACAGAACTCTTGAACAGTAAAGCTGAGCCGCACTGTAACTGAGAGAGATTTG 89895
QY 918 GlyLysAspIleIleAla---GluGluThrProValLeuThrGlnThrLeuProGluGly 936
Db 89896 GAAAGAGATTAATGCAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 89955
QY 937 LysAspAlaHisAspAspMetValThrSerGluValAspPheThrSerGluAlaValThr 956
Db 89956 AGAGAGAGCCCGGGGAGAGAGAGGTCGTTAGTGAAGGAGAAATTAACCCCGAAGCTGAGCA 90015

Qy	957	AlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGlu	976
Db	90016	GCTCAGAAACTGTCAGGGCCATTTGGTGCCGAAGAAGAACCGAAGCATCTGCTGCTGAA	90075
Qy	977	GluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGlu	996
Db	90076	GAGACCACAGAAATGGTGTCAGCAGTCTCCCAAGTTAACCGACTCCCAAGACACACAGAG	90135
Qy	997	GluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluArg	1016
Db	90136	GAGGCCACTCCGGTGCAGAGGTGAAGGTGGCGTACCTGCATAGAAGCAAGAGAGG	90195
Qy	1017	GlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValPro	1036
Db	90196	CGGAATCAAGAGGTCTCTCCAGCAGTCGCAAAAAGTCAAGAGGAATCCAGCTGCCT	90255
Qy	1037	AlaThr-----GlnThrValGlnArgThrGlySerLysAlaLeu	1049
Db	90256	GGCACCGGTGGCCAGAGATGTGCTTCAGCCTGTGCAGAGAGCA-----	90300
Qy	1050	GluLysValGluGluValGluGluAspSerGluValLeuAlaSerGluLysGluLysAsp	1069
Db	90301	---GAGGCAGAAAGACACAGAGCAGCGCTGAAGCGTCGGGTCTCAAGAAAGACGAGAT	90357
Qy	1070	ValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGlu	1089
Db	90358	GTAGTGTGAAGTAGATGTCTCAGGAGCAAAACTGAGCGCTTTTACAAAGGGAAGGTG	90417
Qy	1090	ThrGlyGlnAlaThrProGluSerLeuGlu---ValProGluValThrAlaAspValAsp	1108
Db	90418	GTGGGGCAGACACCCCAAGAAAGCTTTCAAAAAGCTCTCTCAAGTCACAGAGCATAGAG	90477
Qy	1109	His-----ValAlaThrCysGlnVal-----IleLysLeuGlnGln	1120
Db	90478	TCCAGTGAGCTTGTAAACCACTGTCAAGCCGAACACTTAGCTGGGTAAATCAAGAG	90537
Qy	1121	Leu---MetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThr	1139
Db	90538	ATGTTGATGGAACAGGCTATCCCCCTGACTCGGTGGAACCCCTTACAGACAGTGAGACT	90597
Qy	1140	AsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThr	1159
Db	90598	GATGGAACACCCCGTATGACCGACTTTACGCACCCAGCCACCAACCCAGAAAGACGAGAT	90657
Qy	1160	IleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGlu	1179
Db	90658	GTGGAAATCCATGAGGAGAAATGAGTGCATCTGGTACCCAGTCAAGGGGGCACAGAGCA	90717
Qy	1180	GluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGln	1199
Db	90718	GAGCAGTTCTGCACAGAAAGAGAGGCCCTCCAGCA---CCTTCCAGTTTGTGTCTCCAG	90774
Qy	1200	GluGluHisGlyGluGluProGly---ArgAspValLeuGluProThrGlnGlnGluLeu	1218
Db	90775	GAAGAAACTAAAGAACCAATCAAGAGTGAAGACACTCTAGAGCATACAGATAAAGAGGTG	90834
Qy	1219	ThrAlaAlaAlaValProValLeuAlaLysThrGluValGlyGlnGlnGlyValValAsp	1238
Db	90835	TCAGTGGAACTGTATCCATCTGTCAAGAGTGAAGGAGCTCAA-----GAGGCTGAC	90888
Qy	1239	TrpLeuAspGlyGluLysValLys-----GluGluGlnGluValPheVal	1253
Db	90889	CAGTATGCTGATGAGAAACCAAGACGTAACATTTTTCGAAGGACTTGAGGGGTCTATA	90948
Qy	1254	HisSerGly-----ProAsnSerGlnLysAlaAlaAspValThrTyAspSerGluVal	1271
Db	90949	GACACAGCATAAACAGTCAAGTCGGGAAAAGGTCTACTGAAGTTGCGCTTAAAGGTGAAGG	91008
Qy	1272	MetGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSer-----	1287
Db	91009	ACAGAAGAGCTGAATGTAAAGAAAGATGATGCTCTTGAACCTGCAGAGTCACGCTAAGTCT	91068
Qy	1288	---LeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThr	1306

Search completed: December 13, 2002, 03:13:17
Job time : 9259 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

QM protein - nucleic search, using frame_plus_p2n model
Run on: December 13, 2002, 00:16:41 ; Search time 619 Seconds
(without alignments)
5806.445 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSTGEORSEOPAGSDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US0902432/runat.12122002.141451.25790/app.query.fasta_1.1735
-DB=N Geneseq 101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOPC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCAL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0902432 @CNG 1 1 488 @runat.12122002.141451.25790 -NCPU=6 -ICPU=3
-NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8041	99.6	5200	18	RAV02302
2	6001	74.3	5074	18	AAV02301
3	4340.5	53.8	6614	24	ABN59861
4	4300.5	53.3	6608	24	ABN59842
5	4300.5	53.3	6608	24	ABL61938
6	4281	53.0	6605	19	AAV23545
7	4281	53.0	6605	21	AAV74903
8	4226.5	52.4	6886	23	AA585205
9	642	8.0	8307	23	ABL23161
10	639.5	7.9	11597	23	ABL23160
11	565.5	7.0	18506	23	ABL05592
12	557.5	6.9	20448	23	ABL19989
13	537	6.7	24971	23	ABL19988
14	530.5	6.6	24789	23	ABL28640
15	492	6.1	7568	23	AA580823
16	492	6.1	9220	23	ABV24414
17	492	6.1	9220	23	ABV25265
18	492	6.1	9416	24	ABK83800
19	490.5	6.1	10242	23	ABL20787
20	462	5.7	5361	18	AAV78868
21	462	5.7	5529	24	ABK50883
22	462	5.7	6152	18	AAV78867
23	454	5.6	6919	23	ABL05593
24	453.5	5.6	4407	23	ABL12349
25	451	5.6	6561	23	ABL12348
26	450.5	5.6	20978	23	ABL20786
27	448	5.5	9953	23	ABL21371
28	440.5	5.5	5439	23	ABL02411
29	432	5.4	9373	23	ABL07021
30	431.5	5.3	7997	22	ABN47417
31	431.5	5.3	7997	22	ABN47417
32	431.5	5.3	7997	22	ABN47417
33	431.5	5.3	7997	22	ABN47417
34	431.5	5.3	7997	22	ABN47417
35	431.5	5.3	7997	22	ABN47417
36	431.5	5.3	7997	22	ABN47417
37	431.5	5.3	7997	22	ABN47417
38	431.5	5.3	7997	22	ABN47417
39	431.5	5.3	7997	22	ABN47417
40	430	5.3	12505	23	ABL07020
41	429	5.3	8922	23	ABL04053
42	429	5.3	15913	23	ABL04052
43	428.5	5.3	6744	13	AAQ29471
44	427	5.3	15231	24	ABK63598
45	422	5.2	10253	24	ABL60871

ALIGNMENTS

RESULT 1	
AAV02302	
ID	AAV02302 standard; cDNA; 5200 BP.
XX	
XX	
AC	AAV02302;
XX	
DT	20-JUL-1998 (first entry)
XX	
DE	Rat tumour suppressor gene SSeCKS.
XX	
KW	SSeCKS; tumour suppressor gene; rat; protein kinase C; mitosis;
KW	cancer; malignancy; cell proliferation; Alzheimer's disease;
KW	therapy; ss.
XX	
OS	Rattus sp.
XX	
FT	Key Location/Qualifiers
	CDS 1..4791

FT		/**tag= a
XX		
PN	M09740059-AI.	
PD		
XX	30-OCT-1997.	
XX		
PF	18-APR-1997;	97MO-USO6830.
XX		
PR	18-JUN-1996;	96US-0665401.
PR	19-APR-1996;	96US-0635121.
XX		
PA	(GELM/) GELMAN I.	
PA	(JAKE/) JAKEN S.	
XX		
PI	Gelman I, Jaken S;	
XX		
DR	WPI, 1997-535770/49.	
DR	P-PSDB; AAW31347.	
XX		
PT	Tumour suppressor gene SSeCKS - used as a mitotic regulator, and	
PT	inhibitor of malignant phenotype	
XX		
PS	Claim 1; Fig 11A-L; 162pp; English.	
XX		
CC	This rat cDNA sequence codes for the full-length form of the novel	
CC	tumour suppressor protein SSeCKS (see AAW31347). The SSeCKS gene	
CC	product is a substrate of protein kinase C and acts as a negative	
CC	regulator of mitosis and as an inhibitor of the transformed	
CC	phenotype. A partial SSeCKS clone (see AAU02301) was identified by	
CC	searching for cDNAs whose abundance was low in NIH 3T3 cells and	
CC	decreased following the expression of the activated oncogene v-src.	
CC	A 5'RACE product was spliced to this truncated sequence to construct	
CC	a full-length SSeCKS cDNA sequence. This full-length SSeCKS	
CC	nucleic acid sequence, as well as homologous and hybridising nucleic	
CC	acids are claimed, as are vectors comprising such nucleic acids,	
CC	encoded proteins, host cells and methods of inhibiting the	
CC	expression of a transformed phenotype in a host cell by introducing	
CC	the nucleic acid. Introduction of a SSeCKS nucleic acid or gene	
CC	product into a host cell inhibits mitosis of the host cell,	
CC	allowing the treatment of diseases associated with disorders of	
CC	proliferation and/or with the expression of a malignant phenotype.	
CC	SSeCKS can also be used to treat or identify disorders of	
CC	cytoskeletal structure and cellular architecture (such as	
CC	Alzheimer's disease), and may be a marker for aberrancies in	
CC	fertility and/or nervous system development.	
XQ	Sequence 5200 BP; 1612 A; 1201 C; 1563 G; 824 T; 0 other;	
XX		

Alignment Scores:		
Pred. No.:	0	Length: 5200
Score:	8041.00	Matches: 1592
Percent Similarity:	99.75%	Conservative: 0
Best Local Similarity:	99.75%	Mismatches: 4
Query Match:	99.60%	Indels: 0
DB:	18	Gaps: 0
US-09-902-432-4 (1-1596) x AAV02302 (1-5200)		
Oy	1	MetGIyAaGIySeSerThhGIuIaBrgSerProGIuGIaProIaGIySeAaPThr 20
Db	1	ATGGGGCGAGGACAGTTCACCGAGCAGGAGGCCCGAGCAGCGGGGGAGGCACAG 60
Oy	21	ProSeGIuLeuValIeuSeGIyHISGIyProIaAaGIuAaIaSeGIyAaIaGIy 40
Db	61	CCGAGCGAGCTGTCTCAGTGGCCATGGGCCCGCAGCTGAAGCTTCGGAGCAGCTGGA 120
Oy	41	AaSProIaAaSPaIaAaSPProIaThrIySeuProGIuIySaGIyGIuIeuSeSer 60
Db	121	GACCCCGCGAGCGGAGCCCGGCCCAAGAGTCCACAGAGAAATGGCAGCTGTCTTCT 180
Oy	61	ValaGIyValaIaGIuGIaGIyAaPValHISValGIuGIuGIuGIuGIuGIuGIyGIu 80
Db	181	GTCACCGCGTGAAGTGAACAAGGAATCTTCATGTCCAAAGGAAAAACAGAGGGGACG 240

[illegible]

QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGly 460
DB 1321 CAGGAAGCTGAGCTGCTGAGGAGCTGATGAAGACGACAGAGATGTGTCTCTGGAGGA 1380
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGlu 480
DB 1381 GACCAACATCACTGACAGACCTTAAGTCTCTGAAGAGACGCTGCCAAACACCCAGAA 1440
QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySer 500
DB 1441 GGCATTGTCTAGTGGTGGAGATGCTGTCTCTCAGGAAGAATCAAGGTACAGGGAAGT 1500
QY 501 ProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys 520
DB 1501 CCCTTGAAGAAATCTTTCAGTAGCTCAGGCTTAAAGAAGCTGTCTGGGAAGACGAGAAG 1560
QY 521 GlyLysArgGlyGlyGlyAspGluGluProGlyGluPyrGlnHisIleHisThrGlu 540
DB 1561 GGGAAACGAGAGGTGGGGAGACGAGAGCTTGGAGATACCAACATTCACCCGAA 1620
QY 541 SerProGluSerAlaAspGluGlnLysGlyLysSerSerAlaSerSerProGluGluPro 560
DB 1621 TCCCCAGAGAGTCTGATGACGAGAAGGAGAGAGCTCTCGTCTGCCCCGAGGAGCCT 1680
QY 561 GluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
DB 1681 GAGGACACCACTGTCTGGAGAAAGGCCCTCGAAGACACCCAGGATGGGGAAGCTGAG 1740
QY 581 GluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTyrAlaSerPhe 600
DB 1741 GAAGGAACACTTCCGATGGAGAGAAAGAGAGAGAGGATCACTCCCTGGGCATCCTTC 1800
QY 601 LysLysMetValThrProLysLysArgValArgProSerGluSerAspLysGluGlu 620
DB 1801 AAAAGATGTGTACACCAAGAAAGGGTCCGAAGACCTTCTGAGAGTGCACAGGAGGAA 1860
QY 621 GluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
DB 1861 GAGCTGGAGAGGTCAAGACGCGCACCTTGTCTCCACTGATAGCACAGTGTCAAGAAATG 1920
QY 641 GlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArgArgVal 660
DB 1921 CAAGATGAAGTCAAACTGTTGTGAGGAACAAAGCCAGAGAACCAAGACGTAGGGTG 1980
QY 661 AspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArgAlaArg 680
DB 1981 GATACCTCAGTGTCTTGGGAAGCACTGATTGTGTCGGATCATCCAAAGAGAGAGCAAG 2040
QY 681 LysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSerHisArg 700
DB 2041 AAGGCATCCTCTTCAGATGATGAAGGAGGGCCAAAGGACACTGGGAGGGACAGTCAAGA 2100
QY 701 AlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
DB 2101 GCAGAGAGCCAGCAAGAACAAGAAGCCGAACGAGACCTGTCTCTGCGACCCACAG 2160
QY 721 GluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGly 740
DB 2161 GAGCAGGACCAAGCGCAAGGAAGTTCTCACCAGCCAGCGGGAAGCCTTCCGAAGGG 2220
QY 741 GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLys 760
DB 2221 GAAGGTCTCTCCACTTGGGAGTCAATTTAAAGATTTAGTCACTCCCAAGAAAAAATCCAAG 2280
QY 761 SerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThrGluIleGlu 780
DB 2281 TCAAACTGGAAGAGAAAGCCGAAGACTCTAGTGTAGAGCAGTTGTCTCACTGAGATCGAA 2340
QY 781 ProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArgLysLys 800
DB 2341 CCGAGTAGAGAGAAATCTTGGGTCTCCATTAAAGAAATTCATCCCCGACCGCGGAAGAAA 2400
QY 801 ArgAlaAspGlyLysGlnGlnAlaThrValGluAspSerGlyProValGluIleAsn 820

DB 2401 AGGCGACAGCGGGAAGCAAGCAAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAAT 2460
QY 821 GluAspAspProAsnValProAlaValValProLeuSerGlyTyrAsnAlaValGluArg 840
DB 2461 GAGGACGACCCCTAATGTCTCCAGCCGTCGTCTCTGAGTATAATGTCAGTGGAGAG 2520
QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal 860
DB 2521 GAGAAGATGGAAGCCAGGGGAATACGGAGCTGCCCCAGCTGCTGGGGCTGTGTACGTG 2580
QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
DB 2581 TCCGAGAGCTCAGTAAGACTCTGGTCCACACTGTGAGTGTCCAGTCAATGATGGACC 2640
QY 881 ArgAlaValThrSerValGluGluArgSerProSerTrpIleSerAlaSerValThrGlu 900
DB 2641 AGGCGAGTCACCACTGCTGAAGAGCGGTCTCTCTCGTGATATCCGCTCCGTAACAGAA 2700
QY 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
DB 2701 CCTCTTGAACACACACAGCGGGAAGCATGCCACCTGTTGAAGAGGTCACTGAAAAAGAC 2760
QY 921 IleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
DB 2761 ATCATTTGCAGAAAGAACTCTGTGCTCACCCAGACGTTACCAGAGGTAAAGATGCCCAT 2820
QY 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
DB 2821 GACGACATGGTCACCACTGAAGTGGATTTCACCTCAGAGCTGTGACAGCCACAGAGACC 2880
QY 961 SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAsp 980
DB 2881 TCAGAGGCTCTCCGTACTGGAAGATTACCGAAGCATCGGGGGCCGAAGAGACACACAGAC 2940
QY 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
DB 2941 ATGGTGTCCGACGTTTCCAGCTGACTGACTCCCCAGACACCCAGAGGAAGCCACCCCA 3000
QY 1001 ValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAla 1020
DB 3001 GTTCAGAGGTAGAGAGTGGTGTCTAGATACAGAGAAGAGAGGCGCCACAGCGCC 3060
QY 1021 IleLeuGlnAlaValAlaAspLysValLysGluLysSerGlnValProAlaThrGlnThr 1040
DB 3061 ATCCTCCAAGCCGTTCGAGACAAGGTGAAGAGAGTCCCAAGTGTCTGCAACCCAGACT 3120
QY 1041 ValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGlu 1060
DB 3121 GTCGAGAGAACCGGGTCAAAAGCACCTGGAGAGGTTGAGGAGGTAGAGAGGACTCCGAA 3180
QY 1061 ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
DB 3181 GTCTGTGCTTCGAGAGAAAGAGAGGACGTTATGCCGAAAGACCCGTGCGAGGAAGTGA 3240
QY 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
DB 3241 GCTGAGCATCTTCACAGGGCTCTGAGACTTGGACAGCTACTCCAGAGAGCCTTGAAGTT 3300
QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
DB 3301 CCTGAAGTCAAGCCAGATGTAGACCATGTGCCACGTCGCCAGGTTATCAAGCTCCAGCAG 3360
QY 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
DB 3361 CTGATGGAACAGCCCGCTGGCCCTCTGAGTCACTCCGAAACCTTGACAGACAGTGAACAAAT 3420
QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
DB 3421 GGAAGCACTCCCTTAGCAGATTTCAGACACTGCAGATGGGACACAGCAAGATGAACCCATT 3480
QY 1161 AspSerGlnAspSerLysAlaThrAlaValArgGlnSerGlnValThrGluGluGlu 1180

D	b	3481	GACGCCAGACAGTAAAGCCACTGCAGCTGTACAGGACATCACAGTCAACAGAAAG	3540
Q	y	1181	AlaAlaThrAlaGlnIlysgIugIuproSerThrLeuProAsnAsnValProAlaGlnI	1200
D	b	3541	GCGGCTACTGCTCAGAAAGAGGAGCCTTGACACTCACTAATATATGTTCCAGCCAGGAA	3600
Q	y	1201	GluHisGlyGluGluProGlyAlaGAspValLeuGluProThrGlnGlnGluLeuThrAla	1220
D	b	3601	GAACATGGGGAGAACCAAGAGAAAGATGTTCTTGAACCTTACACAGCAAGAGCTTACGTCT	3660
Q	y	1221	AlaAlaValProValLeuAlaIlysThrGluValGlyGlnGlyGlyValAlaAspTyrLeu	1240
D	b	3661	GCAGCCGTGCCCCGTTCTTGCCAGAACCTGAGAGTGGGTCAAGAGGGTGAAGTTGACTGGTTG	3720
Q	y	1241	AspGlyGluIlyValIlysgIugIugInGluValPheValHisSerGlyProAsnSerGln	1260
D	b	3721	GATGGAGAAAAGTCAAAAGAAAGAACAGAGAGTGTTGTTCACCTTGACCCAAACGTGCAA	3780
Q	y	1261	LysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluIly	1280
D	b	3781	AAGCGTGTGATGTGACATATGACAGTGAAGTATGGAGTGGCGGGGTTCAGGAAAAAG	3840
Q	y	1281	GluSerThrGluValGlnSerLeuSerLeuGlyGlyGlyMetGluThrAspValGlu	1300
D	b	3841	GAGACTTCTGAAGTGCAGAGCTTACCTTGAGAGAGGAGAGATGGAAACTGCATCTTGA	3900
Q	y	1301	LysGlyIlyAsrGluThrIlysgIugInValSerGluGlnGlyGluGlnGluThrAla	1320
D	b	3901	AAGGAGAAAGAGGAGCAAAAGCCAGACAGATGATGAAAGATGAGCAGGAAACGCC	3960
Q	y	1321	AlaProGluHisGlyGlyThrTyrGlyIlyProValLeuThrLeuAspMetProSerSer	1340
D	b	3961	GCTCTGTGACATGAGAGCACTTACGGGAAACCGACTCTGACACTTGACATGGCCAGCTCA	4020
Q	y	1341	GluAsrGlyIlyValAlaGlnGlySerLeuGlyGlySerProSerLeuProAspGlnAspIys	1360
D	b	4021	GAGAGGGGAGAGGACCTGGAGAGCCTTGGAGAGAGCCCTTCTCTCCACAGCAACAGACAA	4080
Q	y	1361	AlaGlyCysIleGluValGlnValGlnSerLeuAspThrTyrValThrGlnThrAlaGlu	1380
D	b	4081	GCAGTTCCTCATAGAGGTTCAAGTTCAAAAGCTGTGACACAAACGTCACTCAACACAGAGA	4140
Q	y	1381	AlaValGluIlyValIleGluThrValValIleSerGluThrGlyGluSerProGluCys	1400
D	b	4141	GCTGTGAAAAGGTCATAGAAACCGTGTGATTTCCAGAGCAGGTGAAGATCCAAAGTGT	4200
Q	y	1401	ValGlyAlaHisIleuLeuProAlaGluIlySerSerAlaThrGlyGlyHisTyrThrLeu	1420
D	b	4201	GAGAGTGACCACTTATTCACAGCTGGAAGAGTCTCTGCAACGGGTGGCCACTGACTCTT	4260
Q	y	1421	GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIle	1440
D	b	4261	CAGCATGCAGAGGACAGGTATCCCTGGGGGCTGAGTCTCAGGCAGAAATCCATCCCAATC	4320
Q	y	1441	IleValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyIleSerAla	1460
D	b	4321	ATAGTAACTCTGCTCTCTCGAAAGACACCTTCATCTCGACCTTAAAGAGAAAAATAGCCCA	4380
Q	y	1461	SerGlnAsrGluAsrSerGluGlnGluAspIlyProAspAlaGlyProAspAlaAspGly	1480
D	b	4381	TCCCAAGAGAGCCATTCAGAGAAAGAGAGCAAGCCAGATCTGTGTCTTATGCTAGCGGC	4440
Q	y	1481	LysGluSerThrAlaIleGluIlyValLeuIlyAlaGluProGluIleLeuGluLeuGlu	1500
D	b	4441	AAGGAGATTAACGATGTGAAGAAAGTCTCTCAGGCTGAACTGTAGATCTGTGAAGCTTGAG	4500
Q	y	1501	SerIlySerAsnIlyIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg	1520
D	b	4501	AGTAAAGACCAACAAGATGTGCTGAACGTCAATTCAGACAGCCGTGACACAGTTCCACAGT	4560
Q	y	1521	ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGlnAlaProAlaCysArg	1540
D	b	4561	ACAGAAACAGCCCCGAAATCTCATGCTTATGATTCAACAGCCCAAGGATCTCTTCATGACAGG	4620

Oy	1541	LeuaspSerArgGluProAsnArgCysTrpThrIysMetLysAspAlaLysMetLysHis	1560
Dd	4621	CTTGACAGCGGGAGGCCCAACAATCTGTGACAAAAAAGAAGAACCATGAACAC	4680
Oy	1561	ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrrPalaGln	1580
Dd	4661	CCAATTGCCGACGCCACAGAGAGACTTGCAAGTCtGACCCTTTCTGAGGCATGGCTCAG	4740
Oy	1581	ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys	1596
Dd	4741	CCTGGAAATGCTTGCCGCGCTTGACAGTTGAAAAGCCCGGTGTCAAAG	4788
<hr/>			
RESULT 2			
ID	AAV02301	standard; cDNA; 5074 BP.	
XX	AAV02301;		
AC	AAV02301;		
DT	20-JUL-1998	(first entry)	
XX			
DE	Rat tumour suppressor gene SSeCKS (truncated form).		
XX			
KW	SSeCKS; tumour suppressor gene; rat; protein kinase C; mitosis;		
KM	cancer; malignancy; cell proliferation; Alzheimer's disease;		
XX	therapy; ss.		
OS	Rattus sp.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	176..4216	
FT		/*tag= a	
FT		/transl_except= (pos:189..391, aa:Glu)	
FT		/transl_except= (pos:1661..1663, aa:Lys)	
FT		/transl_except= (pos:1668..1670, aa:Lys)	
FT		/transl_except= (pos:2444..2446, aa:Arg)	
FT		/transl_except= (pos:2528..2530, aa:Glu)	
FT	polyA_signal	2528..2530	
FT		/*tag= b	
FT	polyA_signal	5056..5061	
FT		/*tag= c	
XX			
PN	WO9740059-A1.		
PD	30-OCT-1997.		
XX			
PR	18-APR-1997;	97NO-US06830.	
XX			
PR	18-JUN-1996;	96US-0665401.	
PR	19-APR-1996;	96US-0635121.	
XX			
PA	(GELM/) GELMAN I.		
XX	(JAKE/) JAKEN S.		
PI	Gelman I, Jaken S;		
DR	WPI; 1997-535770/49.		
XX	P-PBDB; AAM31346.		
PT	Tumour suppressor gene SSeCKS - used as a mitotic regulator, and		
XX	inhibitor of malignant phenotype		
PS	Inhibitor of malignant phenotype		
XX			
XX	Example 6; Fig 3A-H; 162pp; English.		
CC	This rat CDNA sequence codes for an active truncated form of the		
CC	novel tumour suppressor protein SSeCKS (see AAM31346). The SSeCKS		
CC	gene product is a substrate of protein kinase C and acts as a		
CC	negative regulator of mitosis and as an inhibitor of the transformed		
CC	phenotype. The SSeCKS clone was identified by searching for cDNAs		
CC	whose abundance was low in NIH 3T3 cells and decreased following the		
CC	expression of the activated oncogene v-src. A 5'RACE product was		
CC	spliced to the truncated sequence to construct a full-length		
CC	SSeCKS cDNA (see AAV02302). This full-length SSeCKS nucleic acid		

SSECKS cDNA (see AAV02302). This full-length SSECKS nucleic acid

CC sequence, as well as homologous and hybridising nucleic acids, are
 CC claimed, as are vectors comprising such nucleic acids, host cells
 CC and methods of inhibiting the expression of a transformed phenotype
 CC in a host cell by introducing such nucleic acids. Introduction of
 CC a SSeCKS nucleic acid or gene product into a host cell inhibits
 CC mitosis of the host cell, allowing the treatment of diseases
 CC associated with disorders of proliferation and/or with the
 CC expression of a malignant phenotype. SSeCKS can also be used to
 CC treat or identify disorders of cytoskeletal structure and cellular
 CC architecture (such as Alzheimer's disease), and may be a marker
 CC for aberrancies in fertility and/or nervous system development.
 XX
 SQ Sequence 5074 BP; 1501 A; 1151 C; 1441 G; 981 T; 0 other;

Alignment Scores:
 Pred. No.: 3,15e-300 Length: 5074
 Score: 6001.00 Matches: 1231
 Percent Similarity: 97.33% Conservatives: 8
 Best Local Similarity: 96.70% Mismatches: 27
 Query Match: 74.33% Indels: 14
 DB: 18 Gaps: 4

US-09-902-432-4 (1-1596) x AAV02301 (1-5074)

QY 329 GluLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArg 348
 Db 2 GAAAGACAGAGCCAGCTCGGAGGAGCAGAGCCGCCAAGAACACAGACCCAGGCCAGG 61
 QY 349 LeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
 Db 62 TTGTCAGCAGACTACGAGAGGTGGAGCTGCCTTTGGAAGCCAGGTTGGTGACCTGGAG 121
 QY 369 AlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
 Db 122 GCATCGTCAGAGGAGAGTGTGCTCTTGGCAAGCAAGTGTGATGAGAGATGGAA 181
 QY 389 AlaHisGlnGluValAlaGluValHisValSerThrValGluLysThrGluGluGlu 408
 Db 182 GCCCACCAGAAAGTTGTGAGAGGTCACGTGAGCACCCTGGAGAACAGACAGAGGAGG 241
 QY 409 GlnGlyGlyGlyGluAlaGluGlyValValValGluGlyThrGlyGluSerLeu 428
 Db 242 CAGGAGGAGGAGGAGGCTGAAGGGGGGCTGGTGTAGAGGAACAGAGAAATCCTTG 301
 QY 429 ProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
 Db 302 CCCCTTGAGAAATCGGCTGAGCCCCAGAGGTCCCCCAGGAAGCTGAGCCTGTGAGGAG 361
 QY 449 LeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeu 468
 Db 362 CTGATGAAGAGCAGAGAGATGTGTGCTCTGGAGGAGACACACTCAACTGACAGACCTA 421
 QY 469 SerProGluGluLysThrLeuProLysHisProGluGlyLeuValSerGluValGluMet 488
 Db 422 AGTCCTGAAGAGAGAGCTGCCCCAACACCCAGAGGCAATTTGACGTAGGTGGAGATG 481
 QY 489 LeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSer 508
 Db 482 CTGTCCTCTCAGGAAGAAATCAAGGTACAGGGAAGTCCCTTTGAAGAAACTCTTCAGTAGC 541
 QY 509 SerGlyLeuLysLysLeuSerGlyLysGlnLysGlnLysArgGlyGlyGlyAsp 528
 Db 542 TCAGGCTTAAAGAGCTGTCTGGGAAGAAGCAGAAAGGGGAAACGAGAGGTGGGGGAGAC 601
 QY 529 GluGluProGlyGluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGln 548
 Db 602 GAAGAGCTCGAGAAATCCAAACACATTCACCCGAATTCCTCAGAGAGTGTGTATGAGCAG 661
 QY 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLys 568
 Db 662 AAGGGAGAGAGCTCTGGTCTGCTCCCGAGAGGCTGAGAGACACAGCTGTCTGGAGAA 721
 QY 569 GlyProLeuGluAlaProGlnAspGlyGluAlaGluGlyThrThrSerAspGlyGlu 588

Db 722 GGCCCGCTGGAAGACACCCAG-GATGGGAAGCTGGAAGAACTACTTCGT--GGAGAG 778
 QY 589 LysLysArgGluGlyIleThrProTyrAlaSerPheLysLysMetValThrProLysLys 608
 Db 779 AAGAAGAGGAAGG--ATCCTCTCTGGGCATCTCTTCAAAAAAGATGGTGACACCCNAGAAA 836
 QY 609 ArgValArgArgProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAla 628
 Db 837 CGGTC-CGAAGACCTTCTGAGAGTCAAGAGGAGAAAGCTGGAGAGGTCGCAAGAGCGCC 895
 QY 629 ThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGly 648
 Db 896 ACCTTGTCTCTCACTGATAGCAGTGTGAGAAATGCAAGATGAAGTCAAACTGTGTGT 955
 QY 649 GluGluGlnLysProGluGluProLysArgValAspThrValSerValSerTyrGluAla 668
 Db 956 GAGGAACAAAAGCCAGAGGAACCAAGCGTAGGCTGGATACTTCAGTGTCTTGGGAAGCA 1015
 QY 669 LeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGlu 688
 Db 1016 CTGATTTGTGTCGATCATCCAAGAAAGAGAGCAAGGAAGGCATCTCTTCAGATATAAGA 1075
 QY 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgAlaGluGluAlaSerLysAspLys 708
 Db 1076 ---GGGCCAAGGACACTGGGAGGGGACAGTCAACAGCAGAGAGAGGCCAGCAAGCAAA 1132
 QY 708 sGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySe 728
 Db 1133 AGAAGCGGA-ACAGAGCGCTGTTCTGCCAGCACCCAGGAGCAGGACCAAGCGCAGGAAG 1191
 QY 728 rSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSe 748
 Db 1192 TTCCTCACCGAGCCAGCGGGAAGCCTTCCGAAGGGGAAGGTGTCTCCACTTTGGGAGTC 1251
 QY 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluLysAla-- 767
 Db 1252 ATTTAAAGATTAGTCACTCCAAGAAAAAATCCAAGTCAAACTGGAAGAGAAGA-AG 1310
 QY 768 ---GluAspSerSerVal---GluGlnLeuSerThrGluIleGluProSerArgGluGl 785
 Db 1311 CCGGAAGGACTCTAGTGTGTAGGAGCAGGTGTCTCACTGAGATCGAACCCGTGTAGAGAGA 1370
 QY 785 uSerTrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLys 805
 Db 1371 ATCTTGGTCTTCCATTAAAGAAATTCATCCCGAGCGCGGAAGAAAGGSCACATGGGAA 1430
 QY 805 s-GlnGlnGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProA 825
 Db 1431 GGCAAGAACCAAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGGACGAGCCTG 1490
 QY 825 snValProAlaValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluA 845
 Db 1491 ATGTCACAGAGCTGTGTCCTCTGTGATGATGATGATGATGATGATGATGATGATGATG 1550
 QY 845 laGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuS 865
 Db 1551 CCCAGGGGAATGCGAGCTGCCAGCTGTGGGGCTGTGTA---GTGTCGAGAGAGCTCA 1607
 QY 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
 Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGACAGGAGGAGTCACCA 1667
 QY 885 erValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHist 905
 Db 1668 GTGTCGAAGAGCGGTCTCTCTGTGATATCCGCTTCCGTAACAGAACCTCTTTGAACACA 1727
 QY 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAlaGluG 925
 Db 1728 CAGCGGGAGAGCCATGCCACCTGTTGAAGAGGTCACTGAAAAAGACATCATTTGCAAGAG 1787
 QY 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945

Db 1788 AAACCTCTGTCTGCACCCAGAGCTTACAGAGGTGCAAGATGCCATGACGATGCTCA 1847
 QY hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
 Db 1848 CCAGTGAAGTGGATTTCACCTCAGAAAGCTGTGACAGCCACAGAGACTTCAGAGGCTCTCC 1907
 QY rgThrGluGluValThrGluAlaSerGluValaGluGluThrThrAspMetValSerAlaV 985
 Db 1908 GTACTGAAGAACTTACCGAAGCATCGGGGGCCGAAGAGACCAAGACATGAGTGTCCGACG 1967
 QY aISerGluLeuThrAspSerProAspThrThrGluGluValaThrProValaGluGluValG 1005
 Db 1968 TTTCCAGCTGACTGACTCTCCAGACACCAAGAGAACCCACAGTTCAGAGAGTGTG 2027
 QY 1uSerGluValaLeuAspThrGluGluGluValaGluGluThrGluAlaIleuGluAlaV 1025
 Db 2028 AGGGTGTGTCTGACTAGATACAGAAAGAGAGAGCCGACAGCAGGAGCCATCTCCAAAGCCG 2087
 QY 1aAlaAspValaValaGluGluGluSerGluValaProAlaThrGluThrValaGluThrThrG 1045
 Db 2088 TTGCAGACAAAGGTGAAAGAGAGTCCAGGTGCTGCAACCAAGACTGTGACAGAAACGG 2147
 QY 1ySerValaLeuGluGluValaGluGluValaGluGluAspSerGluValaIleuAlaSerG 1065
 Db 2148 GGTCAAAAGCACTGAGAGAAAGTTGAGAGAGTGAAGAGGACTCCGAAGTGTGCTTCGG 2207
 QY 1uLySGluValaAspValaMetProLySGluProValaGluGluAlaGluAlaGluAlaV 1085
 Db 2208 AGAAAGAGAGAGCGTTATGCGAAAGAGACCCGTGACAGAAAGCTGGAGCTGAGCATCTTG 2267
 QY 1aGluGluSerGluThrGluGluAlaThrProGluSerLeuGluValaProGluValaThrA 1105
 Db 2268 CACAGAGGCTCTAGAGCTGAGCAGGCTACTCCAGAGAGCCTTGAAAGTCTCTGAAGTACAG 2327
 QY 1aAspValaAspHisValaIaThrCysGluValaIleuLySGluGluGluLeuMetGluGluA 1125
 Db 2328 CAGATGTAGACCATGTGCGCACGTGCGAGTTATCAAGCTCCAGCAGCTGATGGAAACAG 2387
 QY 1aValaIaProGluSerSerGluThrThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
 Db 2388 CCGTGGCCCTGAGTCATCCGAAACCTTGACAGACAGTGAACAATGGAGAGACTCCCT 2447
 QY 1145 euAlaAspSerAspThrAlaAspGlyThrGluGluGluAspGluThrIleAspSerGluAspS 1165
 Db 2448 TAGCAGATTCAACAACCTGAGATGGAGACACAGCAAGATGAAACCATGTGACCGCAGAGACA 2507
 QY eLyAlaIaThrAlaIaValaIaGluGluSerGluValaIaThrGluGluGluAlaIaThrAlaG 1185
 Db 2508 GTAAAGCCACTGACGTGTCAAGGCACTCACAGGTCCAGAAAGAGGCGCTACTGCTC 2567
 QY 1185 1uLySGluGluProSerThrLeuProAsnAsnValaProAlaGluGluGluIleuIleuG 1205
 Db 2568 AAAAAAGAGAGGCTTCGACACTAATTAATGTTCCAGCCCGAAGAAACATGGGGGAG 2627
 QY 1205 1uProGluValaAspValaLeuGluProThrGluGluGluLeuThrAlaAlaIaValaProV 1225
 Db 2628 AACCGAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTGCTGTGAGCGGTGCCG 2687
 QY 1225 aIleuAlaIaThrGluValaGluGluGluGluValaIaAspThrIleuAspGluGluLySv 1245
 Db 2688 TTGGCAAAAGACTGAGGTGGGTCAAGAGGCTGAGTTAATGTTGATGAGAAAG 2747
 QY 1245 aIlySGluGluGluGluValaPheValaHisSerGlyProAsnSerGluValaIaAspV 1265
 Db 2748 TCAAGAAGAGAGAGGAGGTTTGTATCACTTGAGACCAACAGTCAAAAGGCTGCTGATG 2807
 QY 1265 aIThrTyraAspSerGluValaIleuGluValaIaGlyCysGluGluGluSerThrGluV 1285
 Db 2808 TGACATATATGACGTGAAGATGAGGTGGCCGGGTGTAGGAAAGAGAGACTGCTGAAG 2867
 QY 1285 aIGLSerLeuSerLeuGluGluGluGluMetGluThrAspValaGluLySGluValaG 1305
 Db 2868 TGCAAGTCTTTCCTCGAGAGAGAGATGGAACCTGACGTTGAAGAGAGAAAGGG 2927

QY 1305 1uThrLyProGluGluGluValaSerGluGluGluGluGluGluThrAlaIaProGluIleuS 1325
 Db 2928 AGACAAAGCCAGAGCAAGATGAGTGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2987
 QY 1325 1uGluThrTyraGlyLyAspProValaLeuThrLeuAspMetProSerSerGluArgGlyLyS 1345
 Db 2988 AAAGAACTACGGAGAGGCAAGTCTCTGACACTTGACATGCCAGCTCAGAGAGGGGAGAG 3047
 QY 1345 1aLeuGlySerLeuGlyGlySerProSerLeuProAspGluAspLyAlaGlyCysIleG 1365
 Db 3048 CACTGGAAGCCTTGAGAGAGAGCCCTTCTCTCCAGACCAAGACAAAGACAGTGTGATAG 3107
 QY 1365 1uValaGluValaGluSerLeuAspThrThrValaIleuGluThrAlaGluAlaValaGluLySv 1385
 Db 3108 AGGTTCAAGTTCAAGCCTTGACACACAGCATCTCAACAGCAGAGAGAGAGAGAGAGAG 3167
 QY 1385 aIleGluThrValaValaIleSerGluThrGluGluGluSerProGluCysValaGluAlaHis 1405
 Db 3168 TCATGAAACGGTTGTGATTTTCAGAGACAGGTGAAGTCCAGAGTGTAGTGCACACT 3227
 QY 1405 euLeuProAlaGluLySerSerAlaThrGlyGlyHisTyrThrLeuGluHisAlaGluA 1425
 Db 3228 TATTACAGCTGAGAGAGTCTCTGCAACGGGTGGCCACTGAGCTCTTCAGCATGACAGG 3287
 QY 1425 sPThrValaProLeuGluProGluSerGluAlaGluSerIleProIleIleValaThrProA 1445
 Db 3288 ACACGGTACCCCTGGGGCTTCAAGTCTCAGGCAAGATCCATCCCATCTCAATCACTCTG 3347
 QY 1445 1aProGluSerThrLeuHisProAspLeuGluGluGluIleSerAlaSerGluArgGluA 1465
 Db 3348 CTCTGAAAGACCTTAATCTGACTTACAGAGAGAAATTAAGCATCCAGAGAGAGC 3407
 QY 1465 rGSerGluGluGluAspLyAspValaIaGluProAspAlaAspGlyLySGluSerThrA 1485
 Db 3408 GATCAGAGAGAGAGCAAGCAAGTCTGCTGATCTGATGACGAGGCAAGAGAGATACAG 3467
 QY 1485 1aIleGluLyValaLeuLyValaGluProGluIleuGluGluGluSerLySerAsnL 1505
 Db 3468 CAATGCACAAAGTCTCTCAAGCTGAACCTGAGATCTCTGAACTTGAGAGTAAAGACAACA 3527
 QY 1505 ySIIeValaLeuAsnValaIleGluThrAlaValaAspGluPheAlaArgThrGluThrAlaP 1525
 Db 3528 AGATTGTGCTGACGCTCTTACAGACGCCGTTGACCACTTGACAGCTGACAGAAAGGCC 3587
 QY 1525 roGluThrHisAlaTyraAspSerGluThrGluValaProAlaCysArgLeuAspSerArg 1545
 Db 3588 CCGAAACTCATGCTTATGATTCAAGACCAAGGTTCTGCAATGGCTTGAGACAGAGG 3647
 QY 1545 1uProAsnArgCysTyrThrLySmetLyAspAlaLyMetLyHisIaProValaProGluP 1565
 Db 3648 AGCCCAACAGATGCTGGAACAAAGTGAAGATGGCCCAAGATGAACACCCAGGCGCAGC 3707
 QY 1565 roArgGluAspLeuGluValaLeuThrValaLeuGluAlaIaThrAlaGluProArgLySv 1585
 Db 3708 CCAGAGAGAGCTTGCAAGTCTGACCGGTTCTGAGAGCAAGGC--TCAGCTGGAATATCT 3765
 QY 1585 euProArgLeuGluGluGluValaIaProValaSerLyS 1596
 Db 3766 TGCCGCGCTTGAGATTGAAGGCGCGGTGTCAAG 3800

RESULT 3
 ABN59861
 ID ABN59861 standard; cDNA; 6614 BP.
 XX
 AC ABN59861;
 XX
 DT 28-JUN-2002 (first entry)
 XX
 DE Novel human coding sequence SEQ ID NO: 272.
 XX
 KW Human; anti-naemic; vulnerrary; anti-inflammatory; immunomodulator;
 KM antiferility; cerebroprotective; cyostatic; rheumatic; gene therapy;


```

Qy 453 ArgIuMeCysValSerGlyValAspHisThrGluLeuThrAspLeuSerProGluGlu 472
Db 1590 AAAGAAACGTTGTTCCGAGAGAGACCTACACAGAGAGCTGACCTGATGTAGTGG 1649
Qy 473 LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGln 492
Db 1650 AAGGCTCTCCMAACCCCGGAGAGCGCTTGAGTGAAGAGTGGAAATGCTGTCATCAG 1709
Qy 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
Db 1710 GAGAGAAATGAAGTGCAGAGGAAGTCCACTAAAGAGCTTTTATCCAGCCTTGCTTAA 1769
Qy 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyValAspGluGluProGly 532
Db 1770 AAGCTTTCTGGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1826
Qy 533 GluTyrglnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlySer 552
Db 1827 GAGCACACTCAGGTTCCAGCCGATCTCCGACAGCCAGAGAGAGAGAGAGAGAGAGAG 1886
Qy 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1887 TCTGCTCATCCCTCGAGAGAGCCGAGAGATCAAGTCTGAGAAAGGCTTAGCCGAG 1946
Qy 573 AlaProGluAspGlyGluAlaGluGluGlyThrThrSerAspGlyGlyLysLysArgGlu 592
Db 1947 GTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2006
Qy 593 GlyIleThrProThrAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
Db 2007 GGTCGACCTCCCTGGGATCATTCAGAAAGATGCTGACCCCAAGAGAGCTGTAGACGG 2066
Qy 613 ProSerGluSerAspLysGluGluGluGluLysValLysSerAlaThrLeuSerSer 632
Db 2067 CCTTCGAAAGATGATTAAGAGATGAGCTGACAAAGGCTCAAGAGAGAGAGAGAGAG 2126
Qy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlu 652
Db 2127 ACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2186
Qy 653 ProGluGluProLysArgArgValAspThrSerValSerThrProLysLysLeuLysCysVal 672
Db 2187 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2246
Qy 673 GlySerSerLysArgLysArgLysArgLysArgLysArgLysArgLysArgLysArg 692
Db 2247 GGATCATCCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2306
Qy 693 ThrLeuGlyLysAspSerHisArgLysArgLysArgLysArgLysArgLysArgLys 712
Db 2307 GCAATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2366
Qy 713 AspAlaValProAlaSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 732
Db 2367 GACGGGATCTCTGCTGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2426
Qy 733 ProAlaGlySerProSerGluGlyGlyValSerThrThrProLysPheLysArgLeu 752
Db 2427 CAAGCTGGAAGGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2486
Qy 753 ValThrProArgLysLysSerLysSerLysLysLysLysLysLysLysLysLysLys 769
Db 2487 GTACACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2546
Qy 770 ---SerSerValGluGlnLeuSerThrGluGluGluProSerArgGluGluSerThrVal 788
Db 2547 GGGTCTGGTGAACATTCCTCCAGACACTGAAACCGGTTAAAGAGATCTTGGGTC 2606
Qy 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
Db 2607 TCATATCAAAAGATTATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2666
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828

```

```

Db 2667 GCCCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2726
Qy 829 ValValProLeuSerGlyLysArgAlaValGluArgGlyLysMetGlu-----Ala 845
Db 2727 GTGGCTCTGTCTGAGTATGATCTGTAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2786
Qy 846 GlnGlyAsnThrGluLeuProGlnLeuGlnValAlaValThrValSerGluGluLeuSer 865
Db 2787 CAAAAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2846
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2847 GAGATCAGGTTCAATATATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2906
Qy 886 ValGluGluArgSerProSerThrIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2907 ATTGAAGAAAGGCTCTCTTGTGATATCTGCTTCAAGAGAGAGAGAGAGAGAGAGAG 2966
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGlyLysAspIleIleAla---Glu 924
Db 2967 GAAGCTGAAGCCGACCTGTAACTGAGAGAGATTTGAAAGAGAGATTAATGACAGAGA 3026
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3027 GAACCCCAAGGTTACTGAACCTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3086
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3087 GTTAGTGAGCGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3146
Qy 965 ArgThrGluGluValThrGluAlaSerGlyValGluGluGluThrThrAspMetValSerAla 984
Db 3147 GGTTCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3206
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluValAlaThrProValGlnGluVal 1004
Db 3207 GTCTCCAGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3266
Qy 1005 GluSerGlyValLeuAspThrGluGluGluGluGluGluGluGluGluGluGluGlu 1024
Db 3267 GAAGGTGGGCTACGATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3326
Qy 1025 ValAlaAspLysValLysGluGluGluGluGluGluGluGluGluGluGluGluGlu 1038
Db 3327 GTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3386
Qy 1039 ---GlnThrValAlaArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3387 CTTCAGCTGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3428
Qy 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
Db 3429 CAGGCTGAAGCGTCGGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3488
Qy 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
Db 3489 GAGGCAAAAACGACCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3548
Qy 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3549 TTGGAAGAAAGCTCTCAAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3608
Qy 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db 3609 CAAGCGAAACCTTACGCTGGGTAAATATCAACAGAGATGTATGAACAGGCTATCCCC 3668
Qy 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3669 CCTGACTCGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3728
Qy 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167

```

Db 3729 TTTGACCCAGCCAGCAACCCAGAAAGACGAGATTGTGGAAATCCATGAGGAATGAG 3788
Qy 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluAlaAlaThrAlaGlnLysGlu 1187
Db 3789 GTGCGCATCTGTCATCCAGTCAGGGGGGACAGAGACGAGGAGTCTCTGACAGAAAG 3848
Qy 1188 GluProSerThrLeuProAsnValProAlaGlnGluHisGlyGluGluProGly 1207
Db 3849 AGGCTCCAGCA---CCTTCAGTTTGTGTCAGGAAGAACTAAAGAAACAATCAAG 3905
Qy 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu 1226
Db 3906 ATGGAAGACACTAGAGCATACAGATAAGAGGTGTCAGTGGAACCTGTATCCATCTG 3965
Qy 1227 AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys 1246
Db 3966 TCAAAGACTGAGGGGACTCAA-----GAGGCTGACCATGTGATGAGAAACCAA 4019
Qy 1247 -----GluGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db 4020 GAGCTACCATTTTCGAAGAGACTTGAGGGGTCTATAGACACAGGCAATACAGTCAGTCGG 4079
Qy 1260 GlnLysAlaAlaValThrTyAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db 4080 GAAAGGTCACTGAGTTGCCCTTAAGGTGAAGGACAGAGAGCTGAATGTAAAGAG 4139
Qy 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
Db 4140 GATGATGCTCTTGAATGTCAGAGTCAGGCTAAGTCTCTCCATCCCGGTCGAGAGAG 4199
Qy 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
Db 4200 ATGCTAGTTCAAGTTCGAAGGGGAGAAACAGAGACAGACCAACCCATGTGAATGAAGAG 4259
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyGlyLysProValLeu 1333
Db 4260 AAGCTTGAGCAGCAACAGCTGTTACCGTATCTGAAGAGTCAGTAAGAGCTCCTCCAG 4319
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
Db 4320 ACAGTGAATGTGCCATCATAGATGGGCAAGAGAGTTCAGCAGTTCGGAAGAGCCCT 4379
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlyCysAlleGluValGlnValGlnSerLeuAsp 1372
Db 4380 CTTCCCTCCCTAGGTCAGAGGAGGAGTATGCACCAAAATTCAGATTTCAGAGCTCTGAG 4439
Qy 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
Db 4440 GCATCATCTACTTAACAGCGGTCGAGAGGAGGAAAGGTCTTAGGAGAAACTGCCAAC 4499
Qy 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4500 ATTTAGAAACAGGTGAACGTTGGAGCTGCAGTGCACATTTAGTCTTGGAGAGAA 4559
Qy 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4560 TCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGAAGATGCTGTGCCACAGGG 4619
Qy 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
Db 4620 CCCAGTGTGAGGAAATCGACACCATGATGATGATGATGATGATGATGATGATGATGAT 4679
Qy 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluSerGluGluGluAsp 1470
Db 4680 AGTTCCGACCTGGAGAGAGAGAAACACATCACTGAGTGGAGTCAGATGAAGTCGAT 4739
Qy 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaLeuGluLysValLeu 1490
Db 4740 GAGCAGGTTCTGTCAGGAGGAGTC-----AAAGTGAGTGTAGCAATTAGGAGGATTTA 4793
Qy 1491 LysAlaGluProGluLeuLeuGluSerLysSerAsnLysLysLeuValLeuAsnVal 1510
Db 4794 GAGCCTGAAAAATGGGATTTTGGAACTTGAGACCAAGCAAGCAATGTTGTCCAAAACATC 4853

Qy 1511 IleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThrHisAla 1529
Db 4854 ATCCAGACAGCCGTTGACAGGTTTGTACGTACAGAAAGAAACACCCAGCAATGTTGACG 4913
Qy 1530 TyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCys 1549
Db 4914 TCTGAGTTACAGACACACAGCTACGTGATAAAGCTGACAGCCAGGACGCTGGACAGGAA 4973
Qy 1550 TrpThrLys-----MetLysAspAla-LysMetLysHisProValProGlu 1564
Db 4974 ACGGAAAGAGAGGAGGAGGAGCCCTCAGGCTCTGCACAGATGAACACCAATTAATTCA 5033
Qy 1564 nProArgGlu-----AspLeuGlnValLeuThrValLeuGluAl 1577
Db 5034 GCCAAAGAGGAGTCAGAGTCAACCGCAGTGGGACAAAGCACATTTCTGATATTTCCAAAGAC 5093
Qy 1577 aTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5094 ATGAGTGAAGCCCTCAGAAAGAACCATGATGTTGAGGTAGAGGTTCCTCACTG 5145
RESULT 4
ID ABN96942
XX ABN96942 standard; DNA; 6608 BP.
AC ABN96942;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3440 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
WPI; 2002-426119/45.
DR
XX
PT Diagnosing and detecting the progression of liver cancer,
hepatocellular carcinoma or metastatic liver tumor in a patient,
involves detecting the level of expression of two or more genes in a
liver tissue sample
XX
PS Claim 1; SEQ ID NO 3440; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
progression of liver cancer, hepatocellular carcinoma or metastatic liver
tumor in a patient, and differentiating metastatic liver cancer from
hepatocellular carcinoma in a patient, involving detecting the level of
expression of two or more genes represented in ABN93503-ABN97455 in a
tissue sample. The method of the invention has hepatotropic, and
cytostatic activity. The method is useful for diagnosing and detecting
the progression of liver cancer, hepatocellular carcinoma and metastatic
liver carcinoma in a patient. The method is useful for identifying
expression profiles which serve as useful diagnostic markers as well as
markers that can be used to monitor disease states, disease progression,
drug toxicity, drug efficacy and drug metabolism.
CC Note: the sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

Db 2124 ACCGAGACAGCCTCTGAATGCAAGAAATGAAGGAGCGTGGAGAGCCAAAG 2183
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 653 ProGluGluProLysArgArgValAspThrSerValSerTrpGluAlaLeuLeuCysVal 672
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2184 CCGGAGAACCAAGCGCAAGGTGGATACCTCAGTATCTTGGGAAGCTTTAAATTGTGTG 2243
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerSerSerSerSerSerSerSer 692
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2244 GGAATCATCCAGAAAGAGCAAGGAGGAGGTCTCTCTCTGATGAGGAGGAGGACCAAAA 2303
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 693 ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysLysLysLysLysLysLysLys 712
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2304 GCAATGGAGAGACACCAAGAAAGCTGTAGCGCCGGAAGAAAGACAAAGAGACGGGACA 2363
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 713 AspAlaValProLaserThrGlnGlnAspGlnAlaGlnGlnGlnGlnGlnGlnGlnGln 732
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2364 GACGGGATCTTGCTGTTCCCAAGAACATGATCCAGGGCAGGGAAGTTCTCTCCCGGAG 2423
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeu 752
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2424 CAAGCTGGAAGCCCTACCGAAGGGAGGGGTTTCCACCTGGGAGTCATTTAAAGGTTA 2483
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysLysLysLysLysLysLys 769
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2484 GTCACGCCAAGAAATAATCAAGTCCAGCTGGAAGAGAAAGCGAAGACTCCATAGCT 2543
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 770 ---SerSerValGluGlnLeuSerThrGluLeuProSerArgGluGluSerTrpVal 788
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2544 GGGTCTGGTGTAGACATTCCTCCAGACACTGAAACCCGGTAAAGAAAGTCTCTGGGTC 2603
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGln 808
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2604 TCAATCAAGAAGTTTATCTCTGACGAAGCAAGAAAGGCCAGATGGCAACAAAGAACAA 2663
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspProAsnValProAla 828
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2664 GCCCCTGTTGAAGACGCGAGGCCCAACAGGGGCCAACGAAGATGACTCTGATGTCCCGGCC 2723
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2724 GTGGTCCCTCTGCTGAGTATGCTGTAGAAAGGAGAAATGGAGGCACAGCAAGCC 2783
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2784 CAAAAGGCGCAGACGAGCCGAGCAGAGCAGAGCCACTGAGGTGTCCAAGAGCTCAGC 2843
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2844 GAGAGTCAGGTTTCATATGATGGCAGCAGCTGCTGACGGGACGAGGCGCTTACCATT 2903
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2904 ATTGAAGAAAGGTCCTCTCTTGATATCTGTTCACTGACAGAACCTCTTGAAACAAGTA 2963
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 2964 GAAGCTGAAGCGCACCTGTTAACTGAGGAGGTATTGGAAGAGAGTAATTGCAAGAA 3023
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3024 GAACCCCCACCGTTACTGAACCTCTCCACAGAACACAGAGCGCCGGGCGCACCGGTC 3083
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3084 GTTAGTGAGCGGAATTGACCCCGAGAGCTGTGACAGCTGCAGAAATCGAGGCGCCATTG 3143
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3144 GGTTCGGAAGAAAGCAAGCATCTGCTGCTGAAGAGACACAGAAATGGTGTACGCA 3203
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGluAlaThrProValGlnGluVal 1004
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||

Db 3204 GTCTCCAGTTAAACCACTCCCCAGACACCAACAGAGAGGCCACTCGGTGCGAGAGGTG 3263
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3264 GAAGGTGGGTACCTGCATAGAAAGAGAGAGAGAGCGGACTCAAGAGGTCTCTCCAGGCA 3323
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3324 GTGGCAGAAAAGTGAAGAGGAATCCAGCTCGTGGCACCGGTGGGCCAGAGATGTG 3383
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3384 CTTTACGCTGTGCAGAGCA-----GAGGCAGAAAGACCAGAAAG 3425
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1058 AspSerGluValLeuAlaSerGluLysGluLysValMetProLysGlyProValGln 1077
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3426 CAGGCTGAAGCGTCGGTCTGAAGAAAGAGACGGATGTAGTTGAAAGTAGATGCTCAG 3485
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3486 GAGGCAAAACTGAGCCTTTTACACAAGGGAAGGTGGTGGGGCAGACCACCCAGAAAGC 3545
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3546 TTTGAAAAGACTCTCAAGTCACAGAGCATAGATCCAGTGAGCTTTGTAACACTTGT 3605
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3606 CAAGCCGAAACCTTAGCTGGGTAAATACAGGAGATGGTGATGGAACAGGCTATCCCC 3665
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3666 CTTGACTCGGTGGAACCCCTACAGACAGTGAGACTGATGGAAGCACCCCGTAGCCGAC 3725
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3726 TTTGAGCGCACCGGCAACCCAGAAAGACGAGATTGTGGAATTCATGAGGAATGAG 3785
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3786 GTCGCATCTGGTACCAGTCAGCGGGGCACAGAGCAGAGCGAGCTCTGCACAGAAAGAG 3845
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluHisGlyGluGluProGly 1207
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3846 AGCCCTCCAGCA---CCTTCCAGTTTGTCTCCAGGAAGAACTAAAGAAACAATCAAAG 3902
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1208 ---ArgAspValLeuGluProThrGlnGlnGluThrAlaAlaValProValLeu 1226
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3903 ATGGAAAGACACTCTAGAGCATACAGATAAAGAGGTGTCTAGTGGAAACTGTATCCATTCTG 3962
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1227 AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyLysValLys 1246
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 3963 TCAAGACTGAGGGGACTCAA-----GAGGCTGACCAGTATCTGATGAGAAACCCAAA 4016
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1247 -----GluGlnGlnGluValPheValHisSerGly-----ProAsnSer 1259
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 4017 GACGTACCATTTTTTCAAGAGCACTTGAAGGCTCTATAGACACAGGCATACACAGTCAGTCGG 4076
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1260 GlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 4077 GAAAAGGTCACTGAAAGTTCCTTAAAGGTGAAGGGGACAGAAAGCTGAAATGTAAGAG 4136
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 4137 GATGATGCTCTTGAATCTGCAGATCAGCTAAGTCTCTCCATCCCCCTGGGAGAGAGAG 4196
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1295 MetGluThrAspValGlnLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 4197 ATGTTAGTTCAAGTCCGAAGGGAGAAAACAGAGCAGACCAACCCATGTGAATGAAGAG 4256
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeu 1333
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||
Db 4257 AAGCTTGACGACGAAACAGCTGTTACCGTATCTGAAGAGGTGTAAGCAGCTCTCTCCAG 4316
|:::||||| |:::||||| |:::||||| |:::||||| |:::|||||

```

QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlySerLeu 1352
Db 4317 ACAGTGAATGTCCCATCATCATGATGGGCAAGAGAAAGAGATTGGAGAAAGAGCCCT 4376
QY 1353 ProSerLeuProAspGluAspLysAlaGlyCysIleGluValGluValGlnSerLeuAsp 1372
Db 4377 CCTCCCTGCTAGTGTACAGAGAGGATGATATGACCAAAATTCAAGTTGAGAGCTCTGAG 4436
QY 1373 ThrThrValThrGlnThrAlaGluValAlaValGluLysValIle-----GluThrValVal 1390
Db 4437 GCATATTACACTCTAAACAGCGGCTGAGAGAGAGAAAGAGCTTTAGAGAAAGCTGCCAAC 4496
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4497 ATTTAGAAACAGGTGAACCTTTGGAGCTTGACAGGTCACATTTAGTTCTGGAGAGGAAA 4556
QY 1411 SerSerAlaThrGlyGlyHisThrThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4557 TCCCTGAAAAAATGAAAGACTTTGCCGCTCATCCAGGGGAAAGATGCTGTGCCACAGGG 4616
QY 1431 ProGluSerAlaAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4617 CCCGACTGTACAGCAAAATCGACACCGATGATATCTGCTACTACCAAGAAAGCTTAA 4676
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluAsp 1470
Db 4677 AGTTCGCACTGGAAGAGAGAAACACATCATCTGAATGGAAAGTCAGATGAATCGAT 4736
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4737 GAGCAGGTTGCTTGCAGAGAGTC-----AAAGTAGTGTAGCAATTAAGATTAA--- 4787
QY 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4788 -----GACCTGAAATGGAATTTTGGAACTTGAACCAAAAGCACTTAAGTTCCAA 4841
QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr--GluThrAlaProGluThr 1527
Db 4842 AACATCATTCACAGACCCCTTGCCAGTTTGTACGTACAGAAAGAAACACCCAGCGAAAG 4901
QY 1528 HisAlaIAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4902 TTGACCTGTGAAGTTACAGACAGCTCACTGATAAAGCTGACAGCGAGCGCTGGA 4961
QY 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVal 1562
Db 4962 CAGGAAACGAGAAAGAGAGAGAGAAAGAACTCAGAGCTCTGCACAGATGAACACCAATT 5021
QY 1562 IProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5022 ACTTCAGCCAAAGAGAGTACAGATCAACCGACGAGTGGCAAGACATCTCTGATTATTC 5081
QY 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5082 AAAGACATGATGAGCTCAGAAAGAACATGACTGTGAGGTAGAGGTTCCACTG 5139

```

```

XX MO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US10838.
XX 05-JUN-2000; 2000US-208473P.
XX 05-JUN-2000; 2000US-209531P.
XX 18-SEP-2000; 2000US-233133P.
XX 18-SEP-2000; 2000US-233617P.
XX 20-SEP-2000; 2000US-234009P.
XX 20-SEP-2000; 2000US-234034P.
XX 20-SEP-2000; 2000US-234052P.
XX 22-SEP-2000; 2000US-234509P.
XX 22-SEP-2000; 2000US-234567P.
XX 25-SEP-2000; 2000US-234923P.
XX 25-SEP-2000; 2000US-234924P.
XX 25-SEP-2000; 2000US-235077P.
XX 25-SEP-2000; 2000US-235082P.
XX 25-SEP-2000; 2000US-235134P.
XX 25-SEP-2000; 2000US-235280P.
XX 26-SEP-2000; 2000US-235637P.
XX 26-SEP-2000; 2000US-235638P.
XX 27-SEP-2000; 2000US-235711P.
XX 27-SEP-2000; 2000US-235720P.
XX 27-SEP-2000; 2000US-235840P.
XX 27-SEP-2000; 2000US-235863P.
XX 28-SEP-2000; 2000US-236028P.
XX 28-SEP-2000; 2000US-236032P.
XX 28-SEP-2000; 2000US-236033P.
XX 28-SEP-2000; 2000US-236034P.
XX 28-SEP-2000; 2000US-236109P.
XX 28-SEP-2000; 2000US-236111P.
XX 29-SEP-2000; 2000US-236842P.
XX 29-SEP-2000; 2000US-236891P.
XX 02-OCT-2000; 2000US-237172P.
XX 02-OCT-2000; 2000US-237173P.
XX 02-OCT-2000; 2000US-237278P.
XX 02-OCT-2000; 2000US-237294P.
XX 02-OCT-2000; 2000US-237295P.
XX 02-OCT-2000; 2000US-237315P.
XX 03-OCT-2000; 2000US-237425P.
XX 03-OCT-2000; 2000US-237598P.
XX 03-OCT-2000; 2000US-237604P.
XX 03-OCT-2000; 2000US-237608P.
XX 03-OCT-2000; 2000US-237608P.
XX 01-NOV-2000; 2000US-244867P.
XX 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX WPI; 2002-186264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
XX chemical agent to be tested for anti-neoplastic activity, and
XX determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 275; 44p; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a

```

CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 6608 BP; 2091 A; 1418 C; 1827 G; 1272 T; 0 other;

Alignment Scores:

Pred. No.: 1.7e-212 Length: 6608
 Score: 4300.50 Matches: 967
 Percent Similarity: 68.81% Conservative: 189
 Best Local Similarity: 57.56% Mismatches: 408
 Query Match: 53.27% Indels: 116
 DB: 24 Gaps: 36

US-09-902-432-4 (1-1596) x ABL61938 (1-6608)

QY 1 MetGlyAlaGlySerThrGluAlaArgSerProGluGln---ProAlaGlySerAsp 19
 DB 192 ATGGGCGCCGGAGCTCCACCGACGACGCGCGGAGCGCGCGGAGGAGCTCC 251
 QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
 DB 252 ACGCGGCTGAGCCGAGCCGCGGCGCGCGCTCGCGCGGCGCGCGCGCGCGAGCACC 311
 QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
 DB 312 ACCGCGACCCCGCCATCGCTCGGACCCCGCCACCAAGCTCCTACAGAGAAATGGT 371
 QY 57 GlnLeuSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn 76
 DB 372 CAGCTGTCCACCATCAATGGCTAGCTGAGCAAGATGAGCTCAGCTCCAGGAGGGTGAC 431
 QY 77 GlnGluGlyGln-----GluGlu 82
 DB 432 CTAATGGCCAGAAAGAGAGCCCTGAACGGTCAAGAGCCCTTAACACGCGAGGAGAGAA 491
 QY 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
 DB 492 GAAGTCATTGTCAGGAGTTGGACAGAGAGACTCTGAGATGTGAGCGAAGAGATCC 551
 QY 103 ValGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu 122
 DB 552 GATAAAGAGATGGTACTAAGTCAGCGGTGTTCCAGCATCATCAGATGATGGCGAGAG 611
 QY 123 GluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGluMetVal 142
 DB 612 GAGAACCGAAAT---ATCGAACAGATTCTCTTCAGAAAGCAATTTAGAGAGAGCTAAC 668
 QY 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
 DB 669 CAACCCACTGAGTCCAGGCTAATGATATGGATTAAAGAGGTGTTAAGTTTGTGGC 728
 QY 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
 DB 729 TTTAAATTCATGTGTAAGAAAGGATAAGACAGAGAAGCCTGCACACTGTCCAGTACTCA 788
 QY 183 ValLysLysAspGluGlyGluGlyValAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
 DB 789 GTGAAGAAAGATGAAGGGGAGGGAGCA-----CGAGGGGTGGCGACCCAGGAG 839
 QY 203 ProSerValGluThrAlaValGlyLysSerAlaSerLysGluSerGluLeuLysGlnSer 222
 DB 840 CCNAGCCTT-----GGGGTGGAGAGAGAGCATCCAAAGAAAGCGAACCCAAACAATCT 893
 QY 223 ThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIleProLeuGln 242
 DB 894 ACAGAGAAACCCGAGAGACCTTGAAGCGTGAGCAAGCCAGCCAGAGAAATTTCTCCCCCA 953
 QY 243 AlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLysGlnGlu 262

DB 954 GCCGATCTGGCCAGCAGTG---GAGGAATGCAAGAGGAAGGAAGAAACACAGAA 1010
 QY 263 LysGluProThrLysSerProGluSerProSerProValAsnSerGluThrThrSer 282
 DB 1011 AAAGAACCTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTGACAGTGAACAGGATCA 1070
 QY 283 SerPheLysLysPhePheThrHisGlyTyrAlaGlyTyrArgLysLysThrSerPheLys 302
 DB 1071 ACCTTCAAAAAATTTCTCACTCAAGGTGGCGCGGTGGCGCAAAAAGACCACTTTCAGG 1130
 QY 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
 DB 1131 AAGCCGAAGGAGATGAAGTGGAAAGCTTCAGAGAAGAAAGAAACAAAGAGCCAGAAAAA 1190
 QY 323 ValAspGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
 DB 1191 GTAGACACAGAAGACGGAAGGAGGAGGTTCCTCCGAGAAACTGACCGCTCCGAG 1250
 QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
 DB 1251 CAAGCCACCCACAGAGCGCGGAGAAAGTGCCACAGAGCCCGGTTATCAGCTGAATAT 1310
 QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
 DB 1311 GAGAAAGTTGAGCTGCCCTCAGAGGAGCAAGTCAGTGGCTCGCAGGAGACCTTCTGAAGAG 1370
 QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
 DB 1371 AAACCTGCTCCGTGGCGCAGAAAGTGTTCATGAGAAAAATAGAAAGTCCACCAAGAGAG 1430
 QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGly 412
 DB 1431 GTTGTGGCGAAGTCCACGTCAGCAGCGTGGAGAGAGAACCAAGAGAGCAG----- 1481
 QY 413 GlyGluAlaGluGlyGlyValValValGluGlyThrGlyLysSerLeuProGluLys 432
 DB 1482 -----AAAACGAGGTGGAAAGAACACAGCGGTCTGTGCCAGCTGAAGAA 1526
 QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSer 452
 DB 1527 TTGGTTGGAATGATGACAGAACCTCAGGAAGCCGAACCTGCCAAGAGAGCTGGTGAAGCTC 1586
 QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
 DB 1587 AAAGAAACGTGTCTTTCCGAGAGAGACCCCTACACAGAGGAGCTGACCTCAGCTCATGAG 1646
 QY 473 LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGln 492
 DB 1647 AAGGTGCTGTCCAAACCCCGGAGGCGTTGTGAGTGAGGTGGAATGTGTGTATCATCAG 1706
 QY 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
 DB 1707 GAGAGATGAGGTGAGGGAAGTCCACTAAGAAGCTTTTACCAGAGCTGGCTTAAAA 1766
 QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyAspGluGluProGly 532
 DB 1767 AAGCTTTCTGAAAGAAACAGAAAGGGGAAAAAGA---GGAGGAGGAGACAGGAGTAACGG 1823
 QY 533 GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluLysGlyGluSer 552
 DB 1824 GAGCACACTCAGGTTCACCGGATCTCCGAGACGAGGAGGAGGAGGAGGAGGAGGAGC 1883
 QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
 DB 1884 TCTGCTCATCTCCCTGAGAGCCCGGAGAGATCATCGTGTCTGAAAAAGGGCTTAGCCGAG 1943
 QY 573 AlaProGlnAspGlyGluAlaGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
 DB 1944 GTCCAGCAGGATGGGAAGCTGAAGAAGGAGACTACTTCCGATGGAGAGAAAAAAGAGAA 2003
 QY 593 GlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612

Db 2004 GGTGTCATCCCTGGGCATCATTTCAAAAAGATGTCGACGCCCAAGACGGCTTTAGACGG 2063
Qy ProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
Db CCTTCGAAAGATGATTAAGAAAGATGATGATGACCAAGGCTCAAGAGCGCTCTGCTTCTCC 2123
Qy ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGluGluGluLys 652
Db ACCGAGAGACAGCCTCTGAAATGCAAGAAAGAAATGAAGGAGCGGTGAAAGCAAAAG 2183
Qy ProGluGluProLysArgArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
Db CCGGAAAGAACCAACGCGACGAGGTGATCTCAATCTCTGGGAAGCTTTATTTGTGTG 2243
Qy GlysSerSerLysLysArgLysArgLysLysLysLysSerSerSerAspAspGluGluProArg 692
Db GGATCATCCAAAGAAAAGACAGAGAAAGGTCTCTTGTGATGAGAGAAAGGGGACCAAAA 2303
Qy ThrLeuGluGluLysAspSerHisArgLysArgLysLysLysLysLysLysLysLysLys 712
Db GCAAATGGGAGGAGACCAACGAAAGCTGATGAGCGCGGAAAAGACAAAGAGACGGGGACA 2363
Qy AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGlu 732
Db GACGGGATCTCTGCTGCTGCTCCAAAGAACATGATCCAGGGCAGGGAAGTCTCTCCGGAG 2423
Qy ProAlaGlySerProSerGluGluGluGluLysLysLysLysLysLysLysLysLysLys 752
Db CAAGCTGGAAAGCCTTACGAAAGGGAGGGCGTTTCCACTGGGAGATTTAAAGTTA 2483
Qy ValThrProArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 769
Db GTACAGCCCAAGAAAAAACAAGTCCAAAGCTGGAAGAAAAAAGCAACCAATCCTAGCT 2543
Qy 770 ---SerSerValGluGluLeuSerThrGluIleGluProSerArgGluGluSerTrpAl 788
Db GGGTCTGGTGTGAACATTCCTCACTCCAGACACTGAACCCGGTAAAGAAATCTGTGGTTC 2603
Qy SerLysLysLysPheLysLeuProGlyArgArgLysLysLysLysLysLysLysLysLys 808
Db TCAATCAAGAAATTTATCTCTGAGCAAGAAAGAAAGAAAGCCAGATGGAAACAAAGAA 2663
Qy AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db GCGCCCTGTGAAGACGCGCAAGGCGCAAGGCGCCCAAGAGATGACTGTATGCCGCGCC 2723
Qy ValValProLeuSerGluTrpAsnAlaValGluArgGluLysLysLysLysLysLysLys 845
Db GTGGTCCCTCTGTGTAGATGATGCTGTAGAAAGGAAAGGAAATGGAGGACACAGCAAGCC 2783
Qy GlnGlyAsnThrGluLeuProGluLeuGluLysLysLysLysLysLysLysLysLysLys 865
Db CAAAAGGCGAGACGACGCCGAGAGGACGACGACGAGGTGTCCAAGAGGCTCAGC 2843
Qy LysThrLeuValHisThrValSerValAlaValIleAsnGlyThrArgAlaValThrSer 885
Db GAGAGTCAGATTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2844
Qy ValGluGluArgSerProSerTrpLysSerAlaSerValThrGluProLeuGluHisThr 905
Db ATTGAAGAAAGGCTCTCTTGTGATATCTGCTTCACTGATACAGAACCTTTGAACAAGTA 2963
Qy AlaGluGluAlaMetProProValGluGluValThrGluLysAspLysLysLysLysLys 924
Db GAAAGTCGAAGCCGCACTGTAATGAGAGGATATTGGAAGAAAGAAAGTAATTGAGAAAGA 3023
Qy GlnThrProValLeuThrGlnThrLeuProGluGluLysAspAlaHisAspAspMetVal 944
Db GAACCCCAACGATTACTGAACCTCTGCAAGAGAAACAGAGAGCCCGGGCGCAGACGGTTC 3083
Qy ThrSerGluValAspPheThrSerGluValValThrAlaThrGluThrSerGluAlaLeu 964
Db GTTAGTGAAGCGGAATTGACCCCGAAGCTGTGACAGCTGCAAGAACTCGAGGGCCATTG 3143

Qy ArgThrGluGluValThrGluAlaSerGlyValGluGluThrThrAspMetValSerAla 984
Db GGTTCGGAAGAGAAACGAAGCATCTGCTGTGAAGAGACACAGAAATGTGTGACAGA 3144
Qy ValSerGluLeuThrAspSerProAspThrThrGluGluValAlaThrProValGluGluVal 1004
Db GTCTCCAGTTAAACGACTCTCCCAAGACACCAAGAGAGGCACTCCGATCAGAGAGTG 3263
Qy GlysSerGlyValLeuAspThrGluGluGluGluLysArgGlnThrGlnAlaIleLeuGlnAla 1024
Db GAAGGTGCGTACTGACATGAAAGAGACAGAGGCGGACTCAAGAGTCTCCAGCAGA 3323
Qy ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db GTGGCAGAAAAGGTGAAGAGAAATCCAGCTGCTGGCACCGGTGGCCAGAAAGATG 3383
Qy 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluGluGlu 1057
Db CTTCAGCTGTGCAAGAGCA-----GAGCAGAAAGACCAAGAGAG 3425
Qy AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
Db CAGCTGAAGGCTGGGCTGTGAAGAAAGACGAGTGTATGTTGAAGTATGATGCTCAG 3485
Qy GlnAlaGlyAlaGlnHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
Db GAGGCAAAACCTGAGCCTTTTACACAAAGGAGGTGTGTGGGACAGCCCAAGAAAGC 3545
Qy LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db TTTGAAAAGCTCTCAAGTACAGAGAGCATAGAGTCAAGTCAAGTCTTTAACTTGT 3546
Qy 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db CAAGCCGAACCTTACGTGGGTAAATCAACAGAGAGATGGATGAAGACAGATATCCC 3665
Qy ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db CCTGACTGGGTGGAACCCCTTACAGACAGTGAAGTGAAGAGACCCCGTATGCCAC 3725
Qy SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
Db TTTGACGACCAAGGACCAACCAAGAAAGCAAGATTGGGAAATTCATAGAGAGAAATGAG 3785
Qy ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
Db GTGGCATCTGGTAAACCACTCAGGGGCGACAGAGAGAGGCAAGTTCGCAAGAAAGAG 3845
Qy GlnProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly 1207
Db AGGCTTCAGCA---CTTCCAGTTTGTGTTCAGGAAGAAACATTAAGAAACATTAAG 3846
Qy 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu 1226
Db ATGGAAGACACTTGAACATACAGATTAAGAGGAGTGAAGGAAACTGTATCTTCG 3962
Qy AlaLysThrGluValGlyGlnGluGluGluValAspTrpLeuAspGlyGluLysValLys 1246
Db TCAAGACAGAGGGGAGCACTCA-----GAGGCTGACCAAGTATGCGATGAGAAACCAA 3963
Qy 1247 -----GlnGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
Db GACGTACATTTTTCGAGAGCACTTGAGGGGTCTATAGACAGGACATTAACAGTCAAGT 4017
Qy GlnLysAlaAlaAspValThrThrAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
Db GAAAAGTCACTAGAAAGTGGCTTTAAAGTGAAGGAGACAGAAAGAGCTGAATGAAGAA 4077
Qy LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGluGlu 1294
Db GATGATGCTCTTGAACCTGACAGATCACCGTTAAGTCTCTTCATCCCGCTGGAGAGAG 4196

QY 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
DB 2604 TCRAATCAAGAAGTTTATCTCTGGAAGGAAGAAAGGCCAGATGGGAACAAGAACAA 2663
QY 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspProAsnValProAla 828
DB 2664 GCCCTGTTGAAGACGCGAGGCCCAACAGGGCCCAACGAATGACTCTGTATGTCGCCGCC 2723
QY 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
DB 2724 GTGGTCCCTCTGTCTGAGTATGCTGTAGAAAGGAGAAATGGAGGCACAGCAAGCC 2783
QY 846 GluGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuSer 865
DB 2784 CAAAAGCGCAGAGCGCGGACAGAGCGACGCCACTGAGGTGTCCAGAGGCTCAGC 2843
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
DB 2844 GAGAGTCAGGTTTATATGATGGCAGCAGCTGTCTGCTACCGGACGAGCGCAGTACCATT 2903
QY 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
DB 2904 ATTGAAGAAAGGCTCTCTTCTTGATATCTGCTTCAGTGACAGAACCTCTTGAACAAGTA 2963
QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
DB 2964 GAAGCTGAAGCCGACCTGTTAACTGAGGAGGTATTGGAAGAGAAAGTAATTGCAGAGAA 3023
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
DB 3024 GAACCCCGGCTTACTGAACCTCTCCAGAGAACAGAGCGCGCGGCGCACCGGTC 3083
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrThrSerGluAlaLeu 964
DB 3084 GTTAGTGAGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAACTGCGAGGCCATTG 3143
QY 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
DB 3144 GGTTCGAAGAAGAAACCGAAGCATCTGCTGCTGAAGAGACCAACAGAAATGGTGTACGA 3203
QY 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
DB 3204 GTCTCCCGATTACCGACTCCCGAGACACACAGAGAGGCCACTCCGGTGCAGGAGGTG 3263
QY 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
DB 3264 GAAGGTGGCCTACTGTACATAGAAAGACAGACAGAGCGGACTCAAGAGGTCTCTCCAGCA 3323
QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
DB 3324 GTGGCAGAAAAGTGAAGAAGGAATCCAGCTGCTGGCACCGCTGGGCCAGAGATGTG 3383
QY 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
DB 3384 CTTGAGCCTGTGAGAGCA-----GAGCGCAGAAACCCAGAGAG 3425
QY 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
DB 3426 CAGGCTGAACGCTCGGTCTGAAGAAGACACCGATGTAGTTGAAAGTAGATGTCTCAG 3485
QY 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
DB 3486 GAGGCAAAAACCTGAGCCCTTTTACAAAGGGAAGGTGTGTGGCGCAGCACCCCGCAGAAAGC 3545
QY 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
DB 3546 TTTGAAAAAGCTCTCAAGTCAACAGAGCATGATCCAGTCCAGTGTGTAACCACTTGT 3605
QY 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
DB 3606 CAAGCCGAACCTTAGTCTGGGTAAATAATCAGAGAGATGTGTGAACACGCTATCCCC 3665

QY 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
DB 3666 CTTGACTCTGGTGGAAACCCCTACAGACAGTGAAGTGAAGCACCCTCCCTAGCCGAC 3725
QY 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
DB 3726 TTTGACGCCACCGGCACACCCAGAAAGACGAGATTGTGGAAATCCATGAGGAGATGAG 3785
QY 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluAlaAlaThrAlaGlnLysGlu 1187
DB 3786 GTGCATCTGGTA---CCAGTCAGGGGCACAGAAAGCAGAGCAGTCTCTGCACAGAAAGAG 3842
QY 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly 1207
DB 3843 AGCCCTCCAGCA---CCTTCCAGTTTGTGTTCCAGAAAGAACTAAAGAAACAACANAG 3899
QY 1208 ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeu 1226
DB 3900 ATGGAAGACACTCTAGAGCATACAGATAAAGAGGTCTCAGTGGAAACTGTATCCATTCTG 3959
QY 1227 AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLys 1246
DB 3960 TCAAGACACTGAGGGGACTCAA---GAGGCTGACCAAGTATGCTGATGAGAAACCAA 4013
QY 1247 -----GluGlnGlnGluValPheValHisSerGly-----ProAsnSer 1259
DB 4014 GACGTACCATTTTTTCAGAGGACTTGAGGGGTCTATAGACACAGGCATACAGTCAGTCGG 4073
QY 1260 GlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
DB 4074 GAAAGGTCACTGAAGTTCCTTAAAGGTGAAGGACAGAGAACGCTGAATGTAAAGAG 4133
QY 1280 LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyGlu 1294
DB 4134 GATGATGCTCTTGAAGTGCAGAGTCAGCTAAGTCTCTCCATCCCTCCCTGGAGAGAGAG 4193
QY 1295 MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
DB 4194 ATGGTAGTTCAAGTTCGAAAGGGAGAAACAGAAAGCAGAGCAACCCATGTGAATGAAGAG 4253
QY 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeu 1333
DB 4254 AGCTTGACGACGAAACAGCTGTACCGTATCTGAAGAGGTTCAGTAAGCAGCTCCTCCAG 4313
QY 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
DB 4314 ACAGTGAATGTGCCCATCATAGATGGGCAAGGAAGTCAAGAGTTTGAAGAAGCCCT 4373
QY 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
DB 4374 CTTCCCTGCTAGGTCAAGAGGAGGAGTATGCACCAAAATTCAGTTTCAGAGCTCTGAG 4433
QY 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
DB 4434 GCATCAATTCACCTTAACAGCGCTGCAGAGGAGGAAAGGTCTTAGGAGAAACTGCCAAC 4493
QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
DB 4494 ATTTAGAAACAGGTGAACAGTTGGAGCTGCAGGTGCACATTTAGTTCTGGAAAGAA 4553
QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
DB 4554 TCTCTGAAAAAATGAAGACTTTGCGCTCATCCAGGGGAGAGTCTGTGCCACAGGG 4613
QY 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
DB 4614 CCGCACTGTCCAGCAAAATCCAGCCAGTGTATCTCTACTACTACCAAGAAAGGCTTA 4673
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
DB 4674 AGTTCCGACTCTGGAGGAGAAACCAACATCACTGAAGTGAAGTCAGTAAGTCTGAT 4733
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490

```

Db 4734 GAGCAGGTTGCTTCCAGGAGTC-----AAAGTGAAGTGTGCAATTGACGATTTA--- 4784
Oy 1491 LysAlaGluProGlu-----IleuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4785 -----GAGCCGAAATGGATTTTGGAACTTGGAAACCAAAAGCAGTAACTGTCCAA 4838
Oy 1509 AasnValIleGluThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4839 AACATCATCCACAGACGCCCTTGACCGATTGTGACGTAAGAAACACCCAGCCGAAATG 4898
Oy 1528 HisAlaIAspSerSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4899 TTGACGCTGAGTTACAGACACAGCTCAGTATAAAGCTGACAGCCAGACGCTGSA 4958
Oy 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa 1562
Db 4959 CAGGAAACGAGAAAGAAAGAGAGAACCTCAGGCCTCTGCACAGATGAAACACCAATT 5018
Oy 1562 LProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5019 ACTTCAGCCAAAGAGAGTCAAGTCAACCCGACGTGGACAAGACATTCGTATATTTC 5078
Oy 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuAlaProVal 1594
Db 5079 AAAGACATGATGAGACCTCAGAAAGAACCATGACTGTGAGGTAGAGGTTCCACTG 5136

RESULT 7
AAA74903
ID AAA74903 standard; cDNA; 6605 BP.
XX
AC AAA74903;
XX
XX 26-JAN-2001 (first entry)
XX
DE Human gravin coding sequence.
XX
XX Human; gravin; PKA RII binding site; myasthenia gravis;
XX KW kinase anchoring protein; cAMP dependent protein kinase; ss.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 192..5534
XX FT /*tag= a
XX FT /product= "gravin"
XX
XX US6090929-A.
XX
XX PD 18-JUL-2000.
XX
XX PF 19-DEC-1997; 97US-0994570.
XX
XX PR 19-DEC-1996; 96US-0769309.
XX
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX PI KLauck TM, Scott JD, Nauert JB;
XX
XX DR WPI; 2000-523763/47.
XX
XX DR P-PSDB; AAB15380.
XX
XX PT Novel polynucleotides useful for detecting gravin in patients suffering
XX PT from Myasthenia gravis encodes cAMP-dependent protein kinase-binding
XX PT polypeptide and protein kinase C-binding polypeptide of gravin -
XX
XX PS Example 1; Column 23-36; 34bp; English.
XX
XX CC The present sequence is the coding sequence for human gravin. Gravin is
XX CC an A-kinase anchoring protein (AKAP) which is involved in the
XX CC localisation of cAMP dependent protein kinase A (PKA) via interactions
XX CC between the RII binding region and the PKA regulatory subunit RII. Gravin
XX CC is also an antigen found in myasthenia gravis sufferers, and it is

```

```

CC thought that antibodies to it may be useful in modulating the binding of
CC PKA, and thus aid in the treatment of the disease. This sequence was
CC isolated by first screening a human umbilical vein endothelial cell cDNA
CC library with serum from a myasthenia gravis patient, and then searching a
CC human heart cDNA library for sequences resembling the isolated sequence.
CC This was done because the first sequence obtained was shown to be shorter
CC than the full length coding sequence.
XX
SQ Sequence 6605 BP; 2091 A; 1416 C; 1826 G; 1272 T; 0 other;

Alignment Scores:
Pred. No.: 1,72e-211 Length: 6605
Score: 4281.00 Matches: 965
Percent Similarity: 68.69% Conservative: 189
Best Local Similarity: 57.44% Mismatches: 409
Query Match: 53.03% Indels: 117
DB: 21 Gaps: 37

US-09-902-432-4 (1-1596) x AAA74903 (1-6605)
Oy 1 MetGlyAlaGlySerSerThrGlnGlnArgSerProGlnGln---ProAlaGlySerAsp 19
Db 192 ATGGGCGCGGGAGGCTCCACCGAGCGCAGCCCGGAGCAGCCCGGAGGGAGCTCC 251
Oy 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
Db 252 ACGCGCGCTGAGCCCGAGCCCGCGCGCGCCCTCGCGCGAGGCGCGCCAGACACC 311
Oy 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGluLysAsnGly 56
Db 312 ACCGGGGAACCCCGCATGCTGCTCCGACCCCGGCAACCAAGCTCTACAGAAAGATGT 371
Oy 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGln 76
Db 372 CAGCTGTCACCATCAATGCGGTAGCTGAGCAAGATGACTCAGCCTCCAGAGGGTGAC 431
Oy 77 GlnGluGlyGln-----GlnGlu 82
Db 432 CTAAATGGCCGAAAGAGAGCCCTGAAACGCTCAAGAGCCCTTAACAGCCAGAGGAAGA 491
Oy 83 GluValAlaAspGluAspValGlyGlnArgGluSerGluAspValArgLysAspArg 102
Db 492 GAAGCATTTGTACGAGAGTTGAGACAGAGACTCTGAAGATGTAGGGAAGAGACTCC 551
Oy 103 ValGluGluMetAlaIAsnSerThrAlaValGluAspLleThrLysAspGlyGlnGlu 122
Db 552 GATAAAGAGATGGTAACTAAGTCAAGCGGTGTTCAAGCATCACAGATGATGGCAGAG 611
Oy 123 GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetVal 142
Db 612 GAGAACCGAAT--ATCGAACAGATTCCTTCGAAAGCAATTTGAAAGAGCTAAACA 668
Oy 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
Db 669 CAACCCACTGAGTCCCAAGCTATATGATTTAAGAGGTITTAAGTTTGTGGC 728
Oy 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
Db 729 TTTAATTCACTGTGAAAAAGGATTAAGACAGAGACCTGACACTGTCCAGCTACTCACT 788
Oy 183 ValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
Db 789 GTGAAGAAAGATGAAGGGGAGGAGCA-----GCAGGGGCTGGCGACACCAAGAGAC 839
Oy 203 ProSerValGluThrAlaValGlyLysSerAlaSerLysGluSerGluLeuLysGlnSer 222
Db 840 CCCAGCCTT-----GGGCGTGGAGAGCAGCATCCAAAGAAAGCAACCCAAACATCT 893
Oy 223 ThrGluLysGlnGluGlyThrLeuLysGlnGlnGlnSerSerThrGluIleProLeuGln 242
Db 894 ACAGAGAAACCCGAGAGAGACCTGAAGCGTGAGCAAGCAGCAAGAAATTTCTCCCCA 953
Oy 243 AlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGluGluLysGlnGlu 262

```

Db 954 GCCGAATCTGCCAAGCAGTG- -GAGAAATGCAAGAGGAAGAGAGAAACAAGAA 1010
Qy 263 LysGluProThrLysSerProGluSerProSerProValAenSerGluThrThrSer 282
Db 1011 AAGAACCTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTGACCAAGGATCA 1070
Qy 283 SerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSerPheLys 302
Db 1071 ACCPTCAAAAATTTCTTCACTCAAGGTTGGCCCGCTGGCGCAAAAAGACAGCTTTCAGG 1130
Qy 303 LysSerLysGluAspAspLeuThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
Db 1131 AAGCCGAAGGAGGATGAAGTGAAGCTTTCAGAGAGAAAGGAACAGAGCCAGAAAAA 1190
Qy 323 ValAspGluGluGluLysGluLysGluThrGluProAlaSerGluGlu- - - - - 337
Db 1191 GTAGACACAGAAGAAAGCGAAAGCAGAGGTTGCTCTCCGAGAAACTGACCGCTCCGAG 1250
Qy 338 - - - - -GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
Db 1251 CAAGCCACCCACAGGAGCGCGAGAAAGTGCCACAGCGCCCGCTTATCAGCTGAAATAT 1310
Qy 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1311 GAGAAAGTTGAGCTGCCCCCTCAGAGGAGCAAGTCACTGAGTGGCTCGCAGGACCTTCTGAAGAG 1370
Qy 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln- - -Glu 392
Db 1371 AAACCTGCTCGTGGCGACAGAGTGTTCATGAGAAATAGNAGTCCACCAAGAGAG 1430
Qy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGly 412
Db 1431 GTTGTGCGCAAGTCCACGTCAGCACCGCTGAGGAGAGAACCGAAGAGCAG- - - - - 1481
Qy 413 GlyGluAlaGluGlyGlyValValValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1482 - - - - -AAAAAGGAGTGGAAAGAAACACAGCAGGGTCTGTGCCAGCTGGAAGAA 1526
Qy 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
Db 1527 TTGGTTGAATGGATGCAGAACCTCAGNAGCCGAACCTGCCAAGGAGCTGGTGAAGCTC 1586
Qy 453 ArgGluMetCysValSerGlyLysAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1587 AAAAAACGTGTGTTTCCGAGAGGACCTACACAGGAGCTGACCTCAGTCTCTGTATGAG 1646
Qy 473 LysThrLeuProLysHisProGluGlyLysValSerGluValGluMetLeuSerSerGln 492
Db 1647 AAGTGTCTGTCCAAACCCCGAAGCGGCTGTGTAGTGAGGTGGAATGCTGTATCATCAG 1706
Qy 493 GluArgLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
Db 1707 GAGAGATGAAGTGCAGGAGNAGTCCACTAAGNAGCTTTTACCAGACTGGCTTAAAA 1766
Qy 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGly 532
Db 1767 AAGCTTTCTGAAAGAAACAAGAAAGGAAAAA- - -GGAGGAGGAGACGAGGAATCAGGG 1823
Qy 533 GluTyrGlnHisLysHisThrGluSerProGluSerAlaAspGluGlnLysGlyLysSer 552
Db 1824 GAGCACACTCAGGTTCCAGCGCATTTCTCCGACAGCCAGGAGGAGCAAAAAGGCGAGAGC 1883
Qy 553 SerAlaSerSerProGluGluProGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1884 TCTGCTCATCCCCCTGAGGAGCCCGAGAGATACAGTGTCTGGAAAAAGGCTTAGCCGAG 1943
Qy 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1944 GTGCAGCAGGATGGGAAGCTGAAGAAGGAGCTACTTCCGATGGAGAAAAAAGAGAA 2003
Qy 593 GlyIleThrProThrAlaSerPheLysLysMetValThrProLysLysArgValArg 612

Db 2004 GGTGTCACTCCCTCGGCGATCAATCAAAAAGATGGTGAGCCCCCAAGAGCGGTGTAGACGG 2063
Qy 613 ProSerGluSerAspLysGluGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
Db 2064 CCTTCGGAAGTGTATTAAGAAGATGAGTGGACAAGGTCAAGAGCGCTACTTGTCTTCC 2123
Qy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
Db 2124 ACCGAGAGCACAGCTCTGAAATGCAAGAAGAAATGAAAGGGAGCGTGGAAAGACCAAG 2183
Qy 653 ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuLysVal 672
Db 2184 CCGAAGAACCAAGCCGAGGTGGATACCTCAGTATCTTTGGGAAGCTTTAATTTGTG 2243
Qy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArg 692
Db 2244 GGATCATCAAGAAAAAGAGCAAGAGAGGTCTCTTCTGATGAGGAAGGGGACCAAAA 2303
Qy 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2304 GCAATGGAGGAGACCAACCAAGAAAGCTGATGAGCGCGGAAACACAAAGAGACGGGACA 2363
Qy 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGlu 732
Db 2364 GACGGATCTTCTGCTTCCCAAGAACATGATCCAGGGCAGGGAAGTTCTCTCCCGGAG 2423
Qy 733 ProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeu 752
Db 2424 CAAGCTGGAAAGCCCTACCGNAGGGAGGCGTTTCCACCTGGGAGTCATTTAAAGGTTA 2483
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp- - - - - 769
Db 2484 GTCACGCCAAGAAAAAATCAAGTCCAAGTCCAAGTGAAGAGAAAAAGCGAAGACTCATAGCT 2543
Qy 770 - - -SerSerValGluGlnLeuSerThrGluLysGluProSerArgGluGluSerTrpVal 798
Db 2544 GGTCTGGTGTAGAACATTCCTCCTCAGACACTGAAACCCCGTAAAGAAAGAAATCCTGGGTC 2603
Qy 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
Db 2604 TCAATCAAGAGTTTATTCCTGGACCAAGGAGNAGNAGCCAGATGGGAAACAAGAACAA 2663
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2664 GCCCTGTGTAAGACGACGCGGCCAACAGGGGCCAACGAGATCACTCTGTATGTCGCGCC 2723
Qy 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu- - - - -Ala 845
Db 2724 GTGGTCCCTCTGTCTGAGTATGCTGTAGAAAGGGAGAAAAATGGAGGCACAGCAAGCC 2783
Qy 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2784 CAAAAAGGCGCAGAGCAGCGCCGAGAGAGGCGACCTGTAGGTGTCCAAAGGAGCTCAGC 2843
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2844 GAGAGTCAGGTTTCATATGATGGCAGCAGCTGTCTGTCACGGGACGAGGCGAGCTACCAT 2903
Qy 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrCluProLeuGluHisThr 905
Db 2904 ATTGAAGAAGGTCCTCTTCTGGATATCTGCTTCAGTGACAGAACTCTTTGAACAAGTA 2963
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla- - -Glu 924
Db 2964 GAAGCTGAAGCCGCACTGTTAATCTGAGGAGGTATTGGAAGAGAGAGTAATTCAGAGAA 3023
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3024 GAACCCCCACCGTTTACTGAACCTCTGCCAGAGAACAGAGAGGCCCGCGGCGACACGGTC 3083
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3084 GTTAGTGAGCGGAATTTGACCCCGAAGCTGTGACAGCTGCAGAACTGCAGGGCCATTG 3143

XX

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

Homo sapiens.

XX

WO200175067-A2.

PN

11-OCT-2001.

XX

30-MAR-2001; 2001WO-US08631.

XX

31-MAR-2000; 2000US-0540217.

PR

23-AUG-2000; 2000US-0649167.

XX

(HYSE-) HYSEQ INC.

PA

Drmanac RT, Liu C, Tang YT;

PI

WPI; 2001-639362/73.

DR

P-PSDB; ABG21018.

XX

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

Claim 1; SEQ ID No 21009; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 6886 BP; 2142 A; 1483 C; 1908 G; 1352 T; 1 other;

Alignment Scores:

Pred. No.:	1.16e-208	Length:	6886
Score:	4226.50	Matches:	960
Percent Similarity:	68.48%	Conservative:	187
Best Local Similarity:	57.31%	Mismatches:	407
Query Match:	52.35%	Indels:	121
DB:	23	Gaps:	36

US-09-902-432-4 (1-1596) x AAS85205 (1-6886)

QY	1	MetGlyAlaGlySerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp	19
DB	192	ATGGCGCCGGGAGCTCCAGCAGCAGCGAGCCGCGGAGCGCCCGGAGGGAGTCC	251
QY	20	ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla	38
DB	252	ACGCGGCTGAGCCGAGCCGAGCGCGGCGCCCTCGCGCGAGCGCGCCAGACCC	311
QY	39	AlaClyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly	56

DB	312	ACCGGGACCCCGCCATCGCTCGGACCCCGCCACCAAGCTCTACAGAAGATGGT	371
QY	57	GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluAsn	76
DB	372	CAGCTGTCCACCATCAATGGCGTAGCTGAGCAAGATGAGCTCAGCCTCCAGGGGTGAC	431
QY	77	GlnGluGlyGln-----GluGlu	82
DB	432	CTAAATGGCCGCAAGAGAGCCCTGAACGGTCAAGGAGCCCTAAACAGCCAGGAGGA	491
QY	83	GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg	102
DB	492	GAAGTCATTGTCCGGCAGGTGTGATCATAGACTCTGAAGATGTGAGCAGACGACTCC	551
QY	103	ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu	122
DB	552	GATACAGAGATGACTACTAAGTCAGCGGTAGTTTCAGACATCATGATGGCAGGAG	611
QY	123	GluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGluMetVal	142
DB	612	GAGACACCCGATATAATCGAACAGATTCTTCTCAGAAAGCAATTTAGAAAGAGCTAAC	671
QY	143	GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly	162
DB	672	CACCCACTGAGTCCCGAGGCTAATGATATGATTTGAAGAGGTGTTAAGTTTGTGGC	731
QY	163	PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr	182
DB	732	TTTAAATTCACTGTGAAAAAGGATAAGACAGAGAGCGCTGACACTGTCAGCTACTCACT	791
QY	183	ValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu	202
DB	792	GTGAAGAAAGATGAAGGGAGGAGCA-----GCAGGGCTGCGCGACACCCAGGAC	842
QY	203	ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSer	222
DB	843	CCAGGCTT-----GGGCTGGAGAGGAGCATCCAAAGAAAGCAACCCCAACATCT	896
QY	223	ThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIlePro-LeuGlu	242
DB	897	ACAGAAACCCGACAGAGACCTTGAAGCGTGAGCAAAAGCCACCCAGAAATTTCTCCCCA	956
QY	242	nAlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLysGlnGlu	262
DB	957	AGCGGAATCTGGCAAGCAGTG---GAGGAATGCAAGAGAGAGAGAGAGAGAAACAAGA	1013
QY	262	uLysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrSer	282
DB	1014	AAAAGAACCTTAGCAAGTCTCGAAGATCTCCGACTAGTCCCGTGACCGTGAACAGATC	1073
QY	282	rSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSerPheLys	302
DB	1074	AACCTTCAAAAATTTCTTCACTCAAGTTGGCGCGCTGGCGCAAAAGACCAAGTTTCAG	1133
QY	302	sLysSerLysGluAspAspLeuThrAlaGluLysArgLysGluGlnGlnAlaGluLys	322
DB	1134	GAAGCCGAGGAGGATGAAGTGAAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1193
QY	322	sValAspGluGluGluLysThrGluProAlaSerGluGlu-----	337
DB	1194	AGTAGACACAG	1253
QY	338	-----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTy	353
DB	1254	GCAGGCCACCCACAGAGAGCGGCAAGAGTGCACAGAGCCCGCGGTTCAGCTGAATA	1313
QY	353	rGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuLeuAlaSerSerGluGlu	373
DB	1314	TGAGAAAGTTGAGCTCCCTCAGAGGAGCAAGTCAGTGGCTCGCAGGAGACCTTCTGA	1373
QY	373	uLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---	392
DB	1374	GAAACCTGCTCCGTTGGCGACAGAGTGTGTGATGAGAAATAGAAAGTCCACCAAGA	1433

Db 3549 GCTTTGAAAAAGCTCCTCAAGTACAGAGACATAGAGTCCAGTGCAGCTTGTAAACACTT 3608
Qy 1113 ysGlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaVala 1127
Db 3609 GTCAAGCCGAAACCTTAGCTGGGTAAATACACAGGAGATGGTGTGAAACAGGCTATCC 3668
Qy 1127 laProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaA 1147
Db 3669 CCCCTGATCGGTGGAAACCCCTACAGACAGTGCAGACTGATGAAAGCACCCTGAGCCG 3728
Qy 1147 spSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysA 1167
Db 3729 ACTTTGACGACACAGGCACAAACCCAGAAAGACGAGATGTGGAAATCCATGAGGAGATG 3788
Qy 1167 laThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysG 1187
Db 3789 AGTTCGATCTGTGTACCCAGTACGAGGCGCACAGAGGAGGAGGCTTCCTGCACAGAAAG 3848
Qy 1187 luGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProG 1207
Db 3849 AGAGGCTCCAGCA---CCTTCCAGTTTGTGTCTCCAGGAAGAACTTAAAGAAACAATCAA 3905
Qy 1207 ly---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValL 1226
Db 3906 AGATGGAAGACACTCTAGAGCATACAGATAAGAGAGGTGTGAGTGGAAACTGTATCCATTC 3965
Qy 1226 euAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValL 1246
Db 3966 TGTCAAGACTGTGGGACTCAA-----GAGGCTGACCATGTATGCTGAGAAACCA 4019
Qy 1246 ys-----GluGluGlnGluValPheValHisSerGly-----ProAsnS 1259
Db 4020 AAGACGTACCACTTTTTCGAAGGACTTACGGGCTCTATAGACACAGGCATTAACAGTCAGTC 4079
Qy 1259 erGlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnG 1279
Db 4080 GGGAAAGAGTCACTGAAGTTGCCCTTAAAGTGAAGGACAGAGAGCTGAATGTATAAA 4139
Qy 1279 lylsGluSerThrGluValGlnSer-----LeuSerLeuGluGluGlyG 1294
Db 4140 AGGATGATGCTCTTGAAGTGCAGAGTCAAGTCAAGTCTCCATCCCTCGTGAGAGAG 4199
Qy 1294 luMetGluThrAspValGluGlyLysArgGlu---ThrLysProGluGlnValSerG 1313
Db 4200 AGATGTTAGTTCAAGTCCAAAGGAGGAATACCAAGTACAGAGCCCAACCATGTGAATG 4259
Qy 1313 luGluGly---GluGlnGluThrAlaAlaProGluHisGlyGlyThrTyrGlyLysProV 1332
Db 4260 AAGAGAACTTGTAGCGACGAAACAGCTTTACCGTATCTGAAGAGGTCAGTAAGCAGCTCC 4319
Qy 1332 alLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyLys 1352
Db 4320 TCCAGACAGTGAATGTGCCATCATAGATGGGCAAGGAGAGTCAAGCTTTGAAGGAA 4379
Qy 1352 er---ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerL 1371
Db 4380 GCCCTCTCCCTGCTAGTCAAGGAGGAGGAGTATGCACCAAAATTCAGATTCAAGCT 4439
Qy 1371 euAspThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrV 1389
Db 4440 CTGAGGCACTATTCACTCTAACAGCGCTCGCAGGAGGAGGAAAGGCTTAGGAGAACTG 4499
Qy 1389 alValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaG 1409
Db 4500 CCAACATTTTAGAAACAGGTGAACGTTGGAGCGCTGCAGGTGCACATTTAGTTCTGGAAG 4559
Qy 1409 lylsSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProL 1429
Db 4560 AGAAATCTCTGAAAAAATGAAGACTTTTCCGCTCATCCAGGGAAGATGCTGTGCCCA 4619
Qy 1429 euGlyProGluSerGlnAlaGluSerIleProIleLeuValThrProAlaProGluSerT 1449

Db 4620 CAGGGCCCGACTGTCTAGGCAAAATCGACACCACTGATATCTGTCTACTACCAAGAAAG 4679
Qy 1449 hrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluG 1469
Db 4680 GCTTAAGTTCCGACCTTGGAGGAGAGAGAAACACACATCACTGAAGTGAAGTCAAGTGAAG 4739
Qy 1469 luAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysV 1489
Db 4740 TCGATGAGCAGGTTGCTTCCAGGAGGTC-----AAAGTGAAGTGTAGCAATTTGAGGAGG 4793
Qy 1489 alLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsn-LysIleValLeu 1508
Db 4794 ATTGAGGCTGAAATATGGGATTTTGAACTTGAGACCAAAAGCAGTAAAACTTTGTCCAA 4853
Qy 1509 AsnValIleGlnThrAlaValAspGln-PheAlaArgThr---GluThrAlaProGluTh 1527
Db 4854 AACATCATCCAGACAGCCGTTGACCAAGTTTGTACGTACAGAAAGAACAGCCACCAAT 4913
Qy 1527 rHisAlaTyrAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAs 1547
Db 4914 GTTGACGTCTGAGTTACAGACACAA-----GCTTCACGTGATAAAAGCTGACAGCCAGG 4967
Qy 1547 nArgCysTrpThrLysMetLysAspAlaLys-----Me 1558
Db 4968 ACCG---TGGACAGAAACCGGAGAAAGAGGAGGAGGACCTCAGGCCCTCTGCACAGGAT 5024
Qy 1558 tLysHisProValProGlnProArgGlu-----AspLeuGlnVa 1571
Db 5025 GAAACACCAATTACTTCAGCCAAAGAGGAGTCAAGCTCAACCCAGTGGGACAGCACAT 5084
Qy 1571 lLeuThrValLeuGluAlaTrpAlaGlnProArgLys 1583
Db 5085 TCTGATATTTCCAAAGACATGATGTAAGCTCAGAAA 5121
RESULT 9
ABL23161
ID ABL23161 standard; DNA; 8307 BP.
XX
AC ABL23161;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 20956.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 20956; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB16176-AB130511), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (AB057737-AB072072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 8307 BP, 2347 A; 2316 C; 2191 G; 1453 T; 0 other;

Alignment Scores:

Pred. No.:	6,766-24	Length:	8307
Score:	642.00	Matches:	444
Percent Similarity:	34.68%	Conservative:	250
Best Local Similarity:	22.19%	Mismatches:	735
Query Match:	7.95%	Indels:	572
DB:	23	Gaps:	94

US-09-902-432-4 (1-1596) x ABL23161 (1-8307)

QY 2 G1YAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThrPro 21
DB 1648 GGTGCACCTTCACTTCCGAGAA---ACTGAGGAGCAGATGAGGAGCAAGTCTACTGAA 1704
QY 22 SerGluLeuValLeuSerGlyHisGlyProAla-----AlaGluAla 35
DB 1705 GCTCCAACTCTGCGATGACATTGAGCGCGCCAGCCAACTGAACTCTGAGGAAAGCT 1764
QY 36 SerGlyAlaAlaGlyAspProAla---AspAlaAspProAlaThrIleuProGlnIly 54
DB 1765 ACCGGGAGGAGGAGAAATTTGCTTAAGAAACCAACCTGCTGAGAA-----1812
QY 55 AsnGlyGlnLeuSerSerValAsnGlyValAlaGlu-----GlnIlyAspValHisVal 72
DB 1813 -----GCCAGCATTCCTGCTGTGAGGAGGAGATGTTAAGGAAACCAACAGCT 1860
QY 73 GlnGluGlnAsnGlnGluGlnGluGlnValValAspGluAspValGlyGlnArg 92
DB 1861 GGAGAACTTACAGTGAAGGAGC---GATGAGGAGATCGTTAAGGAAACCTCCAGCAGAA 1917
QY 93 GluSerGluAspValArgGluIlyAspArgValGluGlnMetValAlaAsnSerThrAla 112
DB 1918 GAATCCAGTAGTGAAGTAA-----GATGAGCTCACCAAGTCCACACACCT 1965
QY 113 ValGluAspIleThrIlyAspGlyGlnGluGlnIleSerGluIleIleGluGlnIlePro 132
DB 1966 GCCCGAGAACCTTACCTTGTCTGAGAGAGAGATTCAGCAAG-----GAAACCAACACT 2019
QY 133 AlaSerGluAsnAsnVal-----GluGlnMetValGln-----ProAlaGlu 146
DB 2020 GCCGCGAGAACCTTACCTTGTCTGAGAGAGAGATTCAGCAAGCCCCCGCT---2076
QY 147 SerGlnAlaAsnAspValGlyPheIlyIlyValPheIlyPheValGlyPheIlyPheThr 166
DB 2077 GGAGATCTTACGATTCCTGAGAGAGAGAGATC-----GTCAAGGTCAACC 2121
QY 167 ValIlyIlyAspIlyAsnGluIlySerAspThrValGlnLeuLeuThrValIlyIlyAsp 186
DB 2122 ACCCGCGTGGAGAACCTTACGACGTGAAGCGGAGAGAGATCTCAAGGTCCACACCCC 2181
QY 187 GluGlyGluGlyValAlaGluIlyAspValGlyIlyAspHisGlnIlyProSerValGlu 206
DB 2182 GCGTGGAGAA-----TCTACAGTGAAGAGAGAGCAAGCAATCTTAAAGAAAGC 2229
QY 207 ThrAlaValGlyGlu-----SerAlaSerIlyGluSerGluLeuIlyGlnSerThrGlu 224
DB 2230 ACTCTCTCGCGAGAACCAATCACTGAAGGAGAGAGATGTCATCAAGCGCAACACTTCA 2289
QY 225 LysGlnGluGlyThrIleu-----LysGlnGluGlnSerSerThrGluIleProLeu 241
DB 2290 GCTCCCAATTCGACATCGAAGGCGTCAAGAGAACTGAGACACGCCACAGAGTACAGCA 2349

QY 242 GlnAlaGluSerAspGlnAlaAlaGlu-----GluGlnAlaIlyAspGlu---256
DB 2350 GAAGAAGTCGAGAGACTTGCAGCAAAACCAACAGCCCAATTCGCTGAGGAGAGAAACCC 2409
QY 257 -----GlyGlnGluIly-----260
DB 2410 ATTGCGGAGACGCTAATTCAACTGACGATTCAGTGGCGAAGAGAGATTTAAAGGC 2469
QY 261 -----GlnGluIlyGluProThrIlySerProGluSerProSerSerPro 275
DB 2470 ACCATCCACAGACCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2529
QY 276 Val-----AsnSerGluThrThrSerSerPheIlyIlyPhePheThrHis 290
DB 2530 GTTGTGAGAGACATCTTTCGAGCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2589
QY 291 GlyTrpAlaGlyTrpArgIlyIlyThrSerPheIly-----302
DB 2590 GGTGTCCAGATGCTCGAAG 2649
QY 303 -----LysSerIlyGluAspAsp-----308
DB 2650 GACGAG 2709
QY 309 -----LeuGluThr 311
DB 2710 GATGCCACAGACTTGCAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2769
QY 312 AlaGluIlyAspIlyGluGln-----GluAlaGlu-----321
DB 2770 ACCGAG 2829
QY 322 -----LysValAspGluGluGluGluIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 338
DB 2830 ACTAGTAGTCCCGCGCTTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2889
QY 339 GluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspIlyGluIlyValGluLeu 358
DB 2890 AAGAGTAG 2925
QY 359 ProLeuGluAspGlnValGly-----AspLeuGlnAla-----369
DB 2926 CCAATCAATCTGATATTGGGCGACAGCTGTGTGATTCAGATCCGAAGCTACAGAGAGAGAG 2985
QY 370 ---SerSerGluGluIlyIlyCysAla-----ProLeuAla 379
DB 2986 GAGACCACTGATGAGACTGCTACCGATAGCCAGCACTGTTATCTGCTCTGCTGCTCC 3045
QY 380 ThrGluValPheAspGluIlyMetGlnAlaHisGlnIlyValAlaAlaGluValHisVal 399
DB 3046 CAGGAAGTTCGCTCCAGCACTGCTAAAGTTGACAAATCGCAACGACTTGAAACGAGAGAA 3105
QY 400 SerThrValGluIlyThrGluGluGluGlnIly-----410
DB 3106 CCAACCTCCAGCAACAG 3165
QY 411 ---GlyIlyGlyGluAlaGluGlyGlyValValValGluIlyIlyIlyIlyIlyIlyIlyIly 429
DB 3166 CCGCTGATATTCTGCGAG 3207
QY 430 ProGluIlyIlyValAlaGluProGlnIlyValProGlnIlyAlaGluProAlaGlu-----447
DB 3208 ---AAGACCTATGCGCAACAG 3261
QY 448 -----GluLeuMetIlySer 452
DB 3262 TGCAAATGATTCAG 3321
QY 453 ArgGluMetCysValSerGlyIlyAsp-----HisThrGlnLeuThr 466
DB 3322 TTGAGAGAAATGACCGTGGCGAGATCTTCTGATGGTGTGCGACACTTACATTTGGC 3381
QY 467 AspLeuSerProGlu-----GluIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 481

Db 3382 GATGAGAGCTGAAAGTGTGAAAGCATGAGGAATCTACCGCAAGCCGCAATAAG 3441
||| ||| ||| |||:||||| |||
Qy 482 ILeValSerGluValGluMetLeuSerGlnGluArgIleLys----- 496
||| ||| ||| |||:||||| |||
Db 3442 ATTGACGAGGATGTTTCAGAGATCAGCACTGAGAGATTCCTCAAGGATGTCATCATGCCT 3501
||| ||| ||| |||:||||| |||
Qy 497 -----ValGlnGlySerProLeuLysLeuPheSerSerSerGlyLeuLysLysLeu 514
||| ||| ||| |||:||||| |||
Db 3502 ACTGGCATCTGAGCAGCTCTATCTCATGTC----- 3534
||| ||| ||| |||:||||| |||
Qy 515 SerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGlyGluTyr 534
||| ||| ||| |||:||||| |||
Db 3535 -----AAGCCCGATGAAGAG 3549
||| ||| ||| |||:||||| |||
Qy 535 GlnHisIleHisThrGluSerProGluSerAlaAspGluLysGlyLysSerSerAla 554
||| ||| ||| |||:||||| |||
Db 3550 ATCCAACCACTGACCTCGGTTCTGCACAGTTTGATGAA-----TCTACGACAGCT 3600
||| ||| ||| |||:||||| |||
Qy 555 SerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPro 574
||| ||| ||| |||:||||| |||
Db 3601 AAGGTTGATAAGAAACCAATTGATGAATCCGCTGAGGATGAAGAACCAATT----- 3651
||| ||| ||| |||:||||| |||
Qy 575 GlnAspGlyGluAlaGluGlyThrThrSerAspGlyGluLysLysArgGluGlyLe 594
||| ||| ||| |||:||||| |||
Db 3652 -----GGCAATCCGAGGAGGATTGAAACCAATTGATGAATCTGAGGAGGAC----- 3699
||| ||| ||| |||:||||| |||
Qy 595 ThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgProSer 614
||| ||| ||| |||:||||| |||
Db 3700 -----AAGAAACCGGTTGAGAAATCTGCAGAGATAGAAACCCGTT 3741
||| ||| ||| |||:||||| |||
Qy 615 GluSerAspLysGluGluGluLeuLysValLysSerAlaThrLeuSerSerThrAsp 634
||| ||| ||| |||:||||| |||
Db 3742 GAA---GATTCCGAAGAAGAGAGACCATTCGCGACTGTCACTCCAGCGAGTGAGATT 3798
||| ||| ||| |||:||||| |||
Qy 635 SerThrValSerGluMetGlnAspGluValThrValGlyGluGluGlnLysPro--- 653
||| ||| ||| |||:||||| |||
Db 3799 GAAAAGAAATCCAGCCGAGGAGGAAAGAGACTGAGCTGACTTTGGCGGCCCACT 3858
||| ||| ||| |||:||||| |||
Qy 654 GluGluProLysArgArgValAspThrSerValSerTrpGluAlaLeuLysCysValGly 673
||| ||| ||| |||:||||| |||
Db 3859 GAACAGCCAGGCTCAACTCCGCCAGATTGCCAGACTGCTGAGAGAGGATTGAT 3918
||| ||| ||| |||:||||| |||
Qy 674 SerSerLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThr 693
||| ||| ||| |||:||||| |||
Db 3919 GATAAGCTGGCAACCATCTGCTCCGCTTCTGGCAAGATGAGCTCAAGCCT----- 3972
||| ||| ||| |||:||||| |||
Qy 694 LeuGlyGlyAspSerHisArgAlaGluGluAlaSer---LysAspLysGlu-----Ala 710
||| ||| ||| |||:||||| |||
Db 3973 ---GCCGATGAAAGAAACCGCACTGAGACAGCGCAGATCCAGATGCTGAAATCCCTGCT 4029
||| ||| ||| |||:||||| |||
Qy 711 GlyThrAspAlaValProAlaSerThrGln----- 720
||| ||| ||| |||:||||| |||
Db 4030 AGCACCGATGAGCCGAGTCTCGACTGAGTGGCTGCTGATCTCGACAGAAACCA 4089
||| ||| ||| |||:||||| |||
Qy 721 GluGlnAspGlnAlaGlnGlySerSerProGlu-----Pro 733
||| ||| ||| |||:||||| |||
Db 4090 GAGGAGGATTCGACTAAAGCAGCCAGCGCTGAGTCTGATAAGTTCTCGAAGTTCCG 4149
||| ||| ||| |||:||||| |||
Qy 734 AlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeuVal 753
||| ||| ||| |||:||||| |||
Db 4150 ACCAGTCTTCTACTGAAATGAG---ATTGAGGAGTCTGACAAATTCACGACAGTCGCT 4206
||| ||| ||| |||:||||| |||
Qy 754 ThrProArgLysLysSerLysLeuLysGluLysAlaGluAspSerSerValGlu 773
||| ||| ||| |||:||||| |||
Db 4207 CCACCAAAATTTCTGTTCTGAGCAACAGAGCCCACTGCTGAGGAAGATCTCGTC--- 4263
||| ||| ||| |||:||||| |||
Qy 774 GlnLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLys--- 792
||| ||| ||| |||:||||| |||
Db 4264 -----CCTGCAACCTTTGAGCCCATCGAAAGCGAATTTGAAGTTAGCAACCAAGAACCA 4317
||| ||| ||| |||:||||| |||
Qy 793 -----PheIleProGlyArgArgLysLysArg----- 801
||| ||| ||| |||:||||| |||

Db 4318 GCCGTGCAAGGACCACCACTGCCACTCTTGTCTCCGCTCAGCCAGAGAAACCAAGTC 4377
||| ||| ||| |||:||||| |||
Qy 802 -----AlaAsp---GlyLysGlnGluGlnAlaThrValGluAsp 813
||| ||| ||| |||:||||| |||
Db 4378 GATGCAGAGACTTCCACTGAGGCTGATATCAGCACCAACCAAGCGCTGAAGTTGAAAG 4437
||| ||| ||| |||:||||| |||
Qy 814 SerGlyProValGluLeuAsnGluAspProAsnVal----- 826
||| ||| ||| |||:||||| |||
Db 4438 GAAGCTTCTGGTGAACATCCGAGTCGGACATGAATCGATGCTGGTCTCTAGTACT 4497
||| ||| ||| |||:||||| |||
Qy 827 ProAlaValValProLeuSerGluTyrAsnAlaValGluArgGluLys---MetGluAla 845
||| ||| ||| |||:||||| |||
Db 4498 CTTGTTCTCTGTTCTGCCGATGAGGACAGACTCCAGCACTGAGAAAGACTGTGAGGCC 4557
||| ||| ||| |||:||||| |||
Qy 846 -----GlnGlyAsnThrGluLeuPro 852
||| ||| ||| |||:||||| |||
Db 4558 GACGACAAAGTTTACCACCTGTTGCGCCACTTGTGGTGATGAAGAGGAGTCCAACCTTGC 4617
||| ||| ||| |||:||||| |||
Qy 853 GlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHisThrVal 872
||| ||| ||| |||:||||| |||
Db 4618 AAGCTTACCACCAAGATATCTTT-----GAAGAGGAAGCGCTGTTGCGGTTTACCACCGCA 4671
||| ||| ||| |||:||||| |||
Qy 873 SerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArg----- 889
||| ||| ||| |||:||||| |||
Db 4672 GCTCCATCAAGGACGACGCTGAGGACCAAGCTCGAAGTTGAAGAGAAACCGATCGAA 4731
||| ||| ||| |||:||||| |||
Qy 889 ----- 889
||| ||| ||| |||:||||| |||
Db 4732 GATGGACAAAAGCCTATTAGGACGAAACCTCTACACCTACTTTCATCTGAAAATGAGATT 4791
||| ||| ||| |||:||||| |||
Qy 890 -----SerProSerTrpIleSerAlaSerValThrGluProLeuGluHis 904
||| ||| ||| |||:||||| |||
Db 4792 GAGCCTGAATCTGACCGTGCACTACTATTGCTCTTCTAAGGAAGAACCTTCCGAACCA 4851
||| ||| ||| |||:||||| |||
Qy 905 ThrAlaGly-----GluAlaMetProValGluGluValThrGluLys 919
||| ||| ||| |||:||||| |||
Db 4852 TCACCTGGTCTCTACTAAGATGAACCTGCCGAACCACTCCACTGATGCTCTCGATCT 4911
||| ||| ||| |||:||||| |||
Qy 920 AspIleAlaGluGluThrPro-----ValLeuThrGlnThrLeuProGluGly 936
||| ||| ||| |||:||||| |||
Db 4912 GAT--GAGAGAAAGAAACGCTGAGTCGGAAGTGCCACAACTGTTGCACCTGCTGCT 4968
||| ||| ||| |||:||||| |||
Qy 937 LysAspAlaHisAspAspMetValThrSerGluValAspPheThrSerGluAlaValThr 956
||| ||| ||| |||:||||| |||
Db 4969 GAGAAGATTCCAAACAGCAGTACTCTCCGATCAGGAACCACTGCACTTCTGCTCT 5028
||| ||| ||| |||:||||| |||
Qy 957 AlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGlu----- 971
||| ||| ||| |||:||||| |||
Db 5029 GTTGGCAAGCCTGATGAAGATGTAGAGAAAGAACCCAGCACCCGAAATTCCTCCTGACGCT 5088
||| ||| ||| |||:||||| |||
Qy 972 ---AlaSerGlyAlaGlu-----ThrThrAspMetValSerAla----- 984
||| ||| ||| |||:||||| |||
Db 5089 CTTGCATCTGTCGAAGAGGATGAGAATTCCTCTACTGATCAAAATTCATCTGAGGTACCA 5148
||| ||| ||| |||:||||| |||
Qy 985 -----ValSerGlnLeuThr 989
||| ||| ||| |||:||||| |||
Db 5149 GAAAAGAAACCTGAAACTCTGCACAGACGCCAGAAAGAGATATTGTTGGCGCACT 5208
||| ||| ||| |||:||||| |||
Qy 990 AspSerProAspThrThrGluGluAlaThrProValGlnGluVal----- 1004
||| ||| ||| |||:||||| |||
Db 5209 GCAGCTCCAAACGACCTCTGATGAAGTGCCCCCGTCCAGAGGCTGCTGAGGAAGTCTCT 5268
||| ||| ||| |||:||||| |||
Qy 1005 -----GluSerGlyValLeuAspThrGluGlu----- 1013
||| ||| ||| |||:||||| |||
Db 5269 GCTGAGATTCCGCGCCCTCGACTGAACTGGAAATTAACAGCAAGCAAGCAACTGCT 5328
||| ||| ||| |||:||||| |||
Qy 1014 -----GluGluArgGlnThr 1018
||| ||| ||| |||:||||| |||
Db 5329 GCTCCAAGTATTGACCGCAAGAACCTATGTTACTGAAATCGATGAAGAAGCCACCA 5388
||| ||| ||| |||:||||| |||
Qy 1019 GlnAlaIleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThr 1038
||| ||| ||| |||:||||| |||
Db 5389 GTGGCTCTATCTCCGAGAGGATGAGAACCAACTGAGGA----- 5430
||| ||| ||| |||:||||| |||

Oy	1039	GlnThrValGlnAlaGTrHrlySerTyrSAlaLeuGlUylValGlUlnlasp	1058
Db	5431	-----GAGAAAACCGATGGAAACAAGGCCAATCGGAAGAAGCCA	5469
Oy	1059	SerGIuValleuAlaSerGIuLys-----GluTyAspValMetProIysGIProval	1076
Db	5470	TCCGAGGAAGAAGAAAGAGAAACCATTGAAACAAGATGTATCTACTAGAGGACCACTA	5529
Oy	1077	Gln---GluAlaGlyAlaGluHisLeuAlaGInglySerGIuThrclyGlnAlaThrPro	1095
Db	5530	TCTACCGAAGCTTCGTGAAGCTGCATCACCAATCTCAGAGGATTAAGCCCTCGACT	5589
Oy	1096	GluSerLeuGIuValProGIuValThrlAlaAspValAspHisValAlaThrtySerGlnAl	1115
Db	5590	GAGGGA---GAAAGTTGCCGAAMAACCTGAGAGCAAGACAGCTTTCTTGCACTGCCAGGCT	5646
Oy	1116	IleuTyLeuGIuGInGInLeuMetGIuGlnAlaValAlaProGIuSerSerGIuThrLeu---	1134
Db	5647	-----CCAGTTGAGAACAAATTCGGAATCTCCACTGAACTGGCA	5685
Oy	1135	--ThrAspSerGIuThrAsnGlySerThrProLeuAlaAspSerAspThrlAlaSpoly	1153
Db	5686	GCTCAAGATGCGACGAAGGCCACAAGGAGGCTCTGTTGACTCGATGAAATATCTCC	5745
Oy	1154	ThrgInGInAspGIuThrIlleAspSer-----GlnAspSerTyAla	1167
Db	5746	GCTCCGTCGACGAGAAGATTCCTCGTTCTGGCGAGAGAGTCCAGGGCCCCAAGATC	5805
Oy	1168	ThrlAlaAlaValAlaArgInSerGIuValThrGIuGIuGlnAlaAlaThrAlaGInLysGIu	1187
Db	5806	ACAATGCTCTCACTCAAAGCT---GCCGAAGAGAGAACTTAAGACACCGCGGAGAGT	5862
Oy	1188	GluProSerThrLeuProAsnMetValProAlaGInGInGluHisGlyGIuLupProGly	1207
Db	5863	GAACCATCTTCC---ACTGACAAAGTCCAGAAATCGAATACCAAGAGCCCAGAG----	5913
Oy	1208	ArgAspValLeuGIuPProThngInGInGluLeuThrlAlaAlaAlaProValleuAla	1227
Db	5914	-----GACGAGACCAAGGCTGACAGAGACCCCTGAGATCTGTATCCCAAGTGGCT	5961
Oy	1228	LysThrGIuValGIgInGInGluGIuGIuValAspTrpLeuAspDolyGIuLys-----	1244
Db	5962	GATGCGCAATCATTCACCTCGCGCCCCAGTTGCTGGTGTGACATTGAALAAAGACGAACA	6021
Oy	1245	-----ValLysGIuGInGInGluValPheValHisSerGIuProAsnSer	1259
Db	6022	GCCACCACTGCTCGCTCGCGGAAGAGAGAAAGATC-----AMGCCAATATC	6069
Oy	1260	GInLysAlaAlaAspValThrTyAspSerGIuValMetGlyValAlaGlyCysGIuGln	1279
Db	6070	GCTCCAGTGTGTAATAATCTCAACATCC-----GAA	6102
Oy	1280	LysGIuSerThrGIuValGInSerLeuSerLeuGIuGInGlyGluMetGIuThrAspVal	1299
Db	6103	AAGAGAGCCCGGAGATGAGCA-----GAAGTGAATCCGGAACCAAGCCACTCCAGCA	6156
Oy	1300	GluLysGIuLysArg-----GluThrLysProGIuGInValSerGIuGIuGln	1316
Db	6157	GAGTGTGACGCGCCAGCAATGTGATGAATAATGCCCTGCAACAATCGAGCCATCATGATAG	6216
Oy	1317	GInGInThrAlaAlaProGluHisGInGly-----	1326
Db	6217	GCTTCACAGCGCTCCACAAAGAGAAATCCACAACCGTTGCTAGCGCAGCTTCGCT	6276
Oy	1337	-----ThrTyGIuLysProValLeu-----	1333
Db	6277	GCTGTTCACATGATGAATTAAGAGATGTCCACACTACGGAACAGATTCGCGATGAGAG	6336
Oy	1334	ThrLeuAspMetProSerSerGIuArgGlyLysAlaLeuGlySerLeuGIuGIuSerPro	1353
Db	6337	GAAAGGCTGCTCTCTCAGAGACGAACAACAACATCATTAATGATTTTCCACGATTCACCC	6396

OY	1354	SeizleuProaspgrinAaplysaAlaIglyCySIIlegluValGlnValGlnSerLeuaAsphr	1373
Dd	6397	ACTGCCCAAGATGATGTAGAAGCAAGACAAACCAGGGCTCCAGT-----GCACCAC	6450
OY	1374	ThyVal-----ThrGlnThrAlaGlua	1381
Dd	6451	ACTGTTTCATCGCCGACTGCTGATTCTTGCTGCTGATTTAGTACTTCCACAGTGAAGTA	6510
OY	1382	ValGluysVal---IIlegluThrVal-----ValIIeserGluThrGlyGlu	1396
Dd	6511	CCATCTCCTGTTGAATGCACAGCAAGACCATGATGATGATCATGTCTCCACAGCTATTGCT	6570
OY	1397	SerProglucySvalIGlyAlahIsleuLeuProAlagluvsSerSerAlaTheIgLy	1416
Dd	6571	CCTCATACCCGATGTGTGTGGTCCACGCTCACCTCCAGTAGAAGAACAGAGCTCCGCT	6630
OY	1417	HistprThrIeuGlnHisAlaIuaasprThrValProIeuGluProgluSerGln-----	1434
Dd	6631	ACTGTGTCTCTCAGATGCTGTAAGAACT---CCAGTTTCCCAGGCTCCCCAGGATTCT	6687
OY	1435	-----AlaIuSerIIleProIIeIleValThr	1443
Dd	6688	GATAAAGACTCCGAGCTCAGAGGCTCCCCAGATGCCAGAGAAATCCACGTACAAGCCA	6747
OY	1444	ProAlaProglu-----SerThrIeuHisProAspleuGlnGlyIleser	1459
Dd	6748	CCTCTTGACGACAAAGATCCCGACCTGTGGCTCTCCACACTGATGATGATGGGCTTCA	6807
OY	1460	AlaserGlnArgIuaIgsersergIugIuIuaIpylsProAsplalGlyPro---Aspla	1478
Dd	6808	GTACCCGAGCTCCTCTTGATGTAAGACAAAGATCCAAACTACTGGGCTCTCTTGATGAG	6867
OY	1479	AspglyIysgIuserrThrAlaIle-----GluIysVal---LeuIysAlaGlu	1493
Dd	6868	GAGAAAGATTCATCCACAGAGAGCTCCTCTTGATGATGAABAAATTCCTGCTCCAGTTAC	6927
OY	1494	ProglulleuIeuGluIeuGluSerIysSerAsnIysIIeValIeuAsnValIIegInThr	1513
Dd	6928	CCCGTGTTTTCGATGTGGAGCCAAACAGCGAGAG-----CCT	6966
OY	1514	AlaValasprGlnPheAlaAargThrGlu---ThrAlaProgluThrHisAlaTyrasper	1532
Dd	6967	GCTGATCTCGATGTGTGTGTGAAGAGACACCGAACCTGTCATGATGTGGAGACC	7036
OY	1533	GlnThrGlnValProAlaCyArgLeuaspserrArgIuProAsnArgCyetrPrthrIys	1552
Dd	7027	AGCACTGATGAGCTACT-----	7044
OY	1553	MetylAspaIalyMetylshIsProValProGlnProArgIuaSpleuGlnValIeu	1572
Dd	7045	--TCTGATGCGMAATCGAAGCACCACTACTTGACCTGCTGATCCCATCGGA---TCT	7098
OY	1573	ThrValIeuGluIuaITrpAlaGlnProArgIysCySleuProArgIeuGlnIeuLysAla	1592
Dd	7099	CCAGCAGCAGGAGGGGAGATGTGCCAGAAACTGCAGCCCGAGACTTGAAAAGAGAGTG	7158
OY	1593	Pro 1593	
Dd	7159	CCA 7161	
RESULT 10			
ABL23160			
ID	ABL23160	standard; DNA; 11597 BP.	
XX	ABL23160,		
XX			
DT	26-MAR-2002	(first entry)	
XX			
DE	Drosophila melanogaster	genomic polynucleotide SEQ ID NO 20953.	
KW	Drosophila; developmental biology; cell signalling; insecticide;		
KW	pharmaceutical; gene; ds.		
XX			

QY 404 -----LysThrGluGluGluGluGlyGlyGlyGluAlaGluGly 417
 Db 6274 ACTGTGATCTCGACAGAAACCGAGGAGATTCGACTAAAGCCAGCG----- 6327
 QY 418 GlyValValGluGlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGln 437
 Db 6328 -----CTGAGCTCGATTAAGTTCTT----- 6348
 QY 438 GluValProGlnGluAlaGluProAlaGluLeuMetLysSerArgGlyMetCys--- 456
 Db 6349 GAAGTTCGACAGCGCTTCTACTGAAATAGATGAGAGCTCGACAAATTCACAGCA 6408
 QY 457 -----ValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGlu 471
 Db 6409 GTGCTCCACCAAAATTTCTGCTTCTGACGAAACAGACCCT-----GCTGAG 6459
 QY 472 GlyLysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSer 491
 Db 6460 GAAGATCTCGTCCCGAACCTTTGAGCCCATCGAAAGCGAATTTGAGTT-----AGC 6513
 QY 492 GlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeu 511
 Db 6514 ACCAAGAAAGCCGCGTGAAGAGCACACCATCGCCACTTGTCCGCTCAGCCAGAG 6573
 QY 512 LysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyAspGluPro 531
 Db 6574 AAGAAA-----CCAGTCGATCGAGACT 6597
 QY 532 GlyLysLysGlnHisIleHisThrGlu---SerProGluSerAlaAspGluGlnLysGly 550
 Db 6598 TTCACGAGGCTGATATTCAGCCGACGACGAGCCGCTGAAGTTGAAAGAGAGCTTGTGT 6657
 QY 551 GluSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGlnLysGlyPro 570
 Db 6658 GAACATCCGAGTCCGACAAATGA-----ATCATGCTGTGTGT 6696
 QY 571 LeuGluAlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyLysLys 590
 Db 6697 TCTAGTACTCCCTGTTCTGTTCTGTCGATGAGGACAGAACTCCAGCACTGAGAACT 6756
 QY 591 ArgGluGly-----IleThrProTyrAlaSerPheLysMetVal 604
 Db 6757 GTCGAGGCCGACGACAGATTCACCTGTTGCGCCTTGTGTCGATGAGAG----- 6810
 QY 605 ThrProLysLysArgValArgArgProSerGluSerAspLysGluGluLysLeuGlnLys 624
 Db 6811 ---GAGTCCAACTTGGCCCAAGCTACCAAGATATCTTGAAGAGAGAGCGCTGTGG 6867
 QY 625 ValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluVal 644
 Db 6868 GTTACCACCGGCTCCATCAAGGACGCGTGAGCAGAGACGAGTCGAAAGTTGAAGAG 6927
 QY 645 LysThrValGlyGluGluGlnLysProGluGluProLysArgArgValAspHisSerVal 664
 Db 6928 AAACGATGAGAGATGAGAAAGGCTATTGAGAGCAAACTTACACCTTCTATCT 6987
 QY 665 SerTrpGluAlaLeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSer 684
 Db 6988 GAAATATGAG-----ATTGAGCTGATCGACGTCGCAACTATTTGTCTCT 7035
 QY 685 SerAspAspGlyGlyLysProArgThrLeuGlyLys----- 696
 Db 7036 TCTAAGGAAGAACTTCCGAAACCATCCACTGATGCTCTACTAAGAGTAACTCCGAA 7095
 QY 697 AspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValPro 716
 Db 7096 CCATCACTGATGCTCTCTGATCTGATGAGACGAAAGAACCGCTGAGTCCGAAAGTCC 7155
 QY 717 AlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySer 736
 Db 7156 ACAACTGTTCGACCTGCTGCTGAGAAAGATTCCACAAAGAGATATCTCCGATGAGGAA 7215

QY 737 ProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArg 756
 Db 7216 CCA-----ACTGCCACT 7227
 QY 757 LysLysSerLysLysLeuGluGluLysAlaGluAspSerSerValGluLeuSer 776
 Db 7228 TCTGCTCTGTTCGCAACCGCTGATGAATGAGAAAGAAACGACGACGAAATTCCT 7287
 QY 777 ThrGlnIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIleProGly 796
 Db 7288 ACTGACGCTCTGCAATCCGCAAGAG----- 7314
 QY 797 ArgArgLysLysArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSerGlyPro 816
 Db 7315 -----GATGAGATTTCTCTACTGATCAAAATTC 7344
 QY 817 ValGluIleAsnGluAspProAsnValProAlaValProLeuSerGluLysAsn 836
 Db 7345 TCTGAGTACCAAGAAAGAACTGAACTCTCTCAGACGCCA----- 7389
 QY 837 AlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuGly 856
 Db 7390 -----GAAGAGAGATTTGTGGCGCACTGCACAGCTCAACGACCTGTGAT 7437
 QY 857 AlaValTyrValSerGluGluLeuSerLysThrLeuValHisThrVal----- 872
 Db 7438 GAAGTCCCCCGTCCAGAGCTGCTGAGAAAGCTTGTGTGATTCGACGCCCTCG 7497
 QY 873 -----SerAlaIleValIleAspGlyThrArgAlaValThrSerValGluLysArg 889
 Db 7498 ACTGAAATCGAATTAACACAGCAAGCAAACTGCTGCTCCAGATATGACCGCAAG 7557
 QY 890 SerProSerTrpLysSerAlaSerValThrGluProLeuGlnHisThrAlaGlyLysAla 909
 Db 7558 GAACCC-----TATGTACTGAAATCGATGAAGAACCCACC-----ACA 7596
 QY 910 MetProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProValLeu 929
 Db 7597 GTGGCTCTATCTCCGAAAGAGATGAGAACCA---ACTGAGGAAGAAACCACTG--- 7650
 QY 930 ThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSer----- 946
 Db 7651 ---GAACGAAGACCACTGAGAGAGCCATCCAGGAAGAAAGAGAAACCACTT 7707
 QY 947 GluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThr 966
 Db 7708 GAACAAATGATATCTACTGAGGAGCAGTATCTACGAAAGCTTGAAGCTGAGATCTAC 7767
 QY 967 GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaValSer 986
 Db 7768 GAA----- 7770
 QY 987 GlnLeuThrAspSerProAspThrThrGlnGluAlaThrProValGlnGluValGluSer 1006
 Db 7771 -----TCATCTAGGAAGTTAGCCCTCGACTGAG----- 7800
 QY 1007 GlyValLeuAspThrGluGluGluGluLysArgGlnThrGlnAlaIleLeuGlnAlaValAla 1026
 Db 7801 GGAGAAAGTTGCCGAAAGCTCGAGGACGAGCTTCTGACGAGCCAGGCTCCAGTT 7860
 QY 1027 AspLysValLysGlu---GluSerGlnValProAlaThrGlnThrValGlnArgThrGly 1045
 Db 7861 GAGCAATTCGCAAAATCTCCATGACGCTCAAGATGGCGAGAGCCCAAGC 7920
 QY 1046 SerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerGlu 1065
 Db 7921 GAGGCTCTGTTGACTCC-----GATGAAGTATCTCCGCTCCGTCGACAG 7968
 QY 1066 LysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAla 1085
 Db 7969 AAGATTCCCTCCGTT-----TCTGGGAGAGAA-----GTC 7998
 QY 1086 GlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluVal----- 1103

Db	7999	GAAGGCCCAAGTCAACAATGCTCTCACTCAAGCTGCGCAAGAGGAGCAACTTAAGACA	8058
Qy	1104	-----ThrAlaaspValaspHisValAlaThrCysGlnValIleLeu	1118
Db	8059	CCCGCGAGAGTGAACCATCTTCCACTGACAGGTACCAAGAACTGATACCAAGACCC	8118
Qy	1119	GlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGlu	1138
Db	8119	GAGACGAGACCAAGGTGACGAGACCCCTGATCTGTTCACCAAGTCTGTGATGGCA	8178
Qy	1139	ThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGlu	1158
Db	8179	ACATCCACTTCGCGCCCGATGCTGTGTGTGACATT-----	8214
Qy	1159	ThrIleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGlu	1178
Db	8215	-----GAAGAAGACCAAGCAACCACTGCC-----TCGCTGAAGAGAG	8256
Qy	1179	GluGlu-----AlaAlaThrAlaGlnLysGluGluProSerThrLeuPro	1193
Db	8257	GAAGAGATCAAGCAACTATCGCTCCAGCTGTGTAATTCCTCAACCATCC-----	8307
Qy	1194	AsnAsnValProAlaGlnGluHisGlyGluGlu-ProGlyArgAspValLeuGluPr	1213
Db	8308	GAAGAAGGAGCCGTGGATGACGAAGAAGTGGAAATCCGGAACCAAGG-----CC	8355
Qy	1213	oThrGlnGlnGluLeuThrAlaAlaAlaValProValLeuAlaLysThrGluValGlyG	1233
Db	8356	ACTCCAGCAGAGTCTGACGCCAGCCAA-----TTGATGAATGGCCCTGCA-AC	8405
Qy	1233	ngluGlyGluValAspTrpLeuAspGlyGluLysValLysGluGlnGluValAlaPheVa	1253
Db	8406	ATCTGGACCATCGATGAGCTTCCACAGCCGCTCCCAAAAGAGGAATCCACAACCGT	8465
Qy	1253	LHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTrpAspSerGluValMetG	1273
Db	8466	TGCTAGCGCA-----GCTTCGCCTGCTGTTCACGATGATGAATTAAGGA	8510
Qy	1273	yValAlaGlyCysGln-----GluLysGluSerThrGluValGlnSerLeuSe	1289
Db	8511	TGTCACCACTACGCAACCGATGGCCGATGAGAAGAGAGTGGCTGCTCTCTAG-----	8562
Qy	1289	rLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProG	1309
Db	8563	-----GACGAAACGAACATCCATTGATGTTCCACCGAT-----TCACCCAC	8606
Qy	1309	uGlnValSerGluGluGlyGluGlnGluThrAlaAlaProGluHisGluGlyThrTyrg	1329
Db	8607	TGCCCAAGATGATGAGAAGCAACAGACAAACCGAGCTCCAGTGGCACCACACTGTTTC	8666
Qy	1329	yLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLe	1349
Db	8667	ATCGCCG-----ACTGCTGATTCTGCTGCTGATCT-----	8697
Qy	1349	uGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValG	1369
Db	8698	-----AGTACTCCCACTGGAGGTACCATCTCTCTGTGTAATTCGACACGAA	8744
Qy	1369	nSerLeuAspThrThrValThrGlnThrAlaAlaValGluLysValIleGluThrVa	1389
Db	8745	GCCAATGGATGATCATCATGTGCCAGACT-----	8772
Qy	1389	lValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaG	1409
Db	8773	-ATTGCTCTCATACCGCCGAT-----GGTGTGCGTCCAGTCCACTGA	8816
Qy	1409	uLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLe	1429
Db	8817	AGATGAAGACAGGCTCCCGTTACTGTGTCCTCAGGATGCTGAAAGACT---CCAGT	8873
Qy	1429	uGlyProGluSerGln-----AlaG	1436

Db	8874	TTCCCCGGCTCCCGCAGGATTCTGATTAAGACTCCGAGCTCAGAGGCTCCCGCAGGATGCCGA	8993
Qy	1436	uSerIleProIleIleValThrProAlaProGlu-----SerThrLeuHisPr	1452
Db	8934	CGAATCCGAGCTACACGCACCTCTTCACGACAAACAGATCCCGAGCATCTGGGTCC	8993
Qy	1452	oAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAspLysPr	1472
Db	8994	TCAGACTCATGATGGGGTTCAGCTACCCGAGCTCTCTTGATGAAGACAAGATCCCAAC	9053
Qy	1472	oAspAlaGlyPro--AspAlaAspGlyLysGluSerThrAlaIle-----G	1487
Db	9054	TACTGGCGGCTCTCTTGATGAGGAGAAGATTCATCCACAGCAGCTCTCTTGATGATGA	9113
Qy	1487	uLysVal---LeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIl	1506
Db	9114	AAAGATTCTGCTCCAGTTAGCCCGTGTTCATGTGGAGCCACGACGGAGAAG--	9171
Qy	1506	eValLeuAsnValIleClnThrAlaValAspGlnPheAlaArgThrGlu---ThrAlaPr	1525
Db	9172	-----CCTGCTGTATCTCAGTATGATGTGTGAAGAGACACCGAAC	9212
Qy	1525	oGluThrHisAlaTyAspSerClnThrGlnValProAlaCysArgLeuAspSerArgGl	1545
Db	9213	ACCTGTCATGATGTGAGACACGACACTGATGAGCTACT-----	9252
Qy	1545	uProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHisProValProGlnPr	1565
Db	9253	-----TCTGATGCCAAACTGAAGCCACTACTTCTTGCACC	9287
Qy	1565	oArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCysLe	1585
Db	9288	TGCTACGCCATCGAA--TCTCCAGCCACGGAGCGGAGATGTGTCCAGAAACTGCAGC	9344
Qy	1585	uProArgLeuGlnLeuLysAlaPro	1593
Db	9345	CCCAGACTTGAAGAGGAGGTGCCA	9369
RESULT 11			
ABL05592			
ID ABL05592 standard; cDNA; 18506 BP.			
XX	AC	ABL05592;	
XX	DT	26-MAR-2002 (first entry)	
XX	DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 11258.	
XX	KW	Drosophila; developmental biology; cell signalling; insecticide;	
XX	KW	pharmaceutical; gene; ss.	
XX	OS	Drosophila melanogaster.	
XX	FN	WO200171042-A2.	
XX	PD	27-SEP-2001.	
XX	PF	23-MAR-2001; 2001WO-US09231.	
XX	PR	23-MAR-2000; 2000US-191637P.	
XX	PR	11-JUL-2000; 2000US-0614150.	
XX	PA	(PEKE) PE CORP NY.	
XX	PI	Venter JC, Adams M, Li PWD, Myers EW;	
XX	DR	WPI; 2001-656860/75.	
XX	DR	P-PSDB; ABB61489.	
XX	PT	New isolated nucleic acid detection reagent for detecting 1000 or more	
XX	PT	genes from Drosophila and for elucidating cell signalling and cell-cell	
XX	PT	interactions -	

PS Claim 1; SEQ ID NO 11258; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB161840-AB16175) and the encoded proteins
 CC (AB161737-AB162072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPD
 CC at ftp.wipo.int/pub/published_pcf_sequences.

XX Sequence 18506 BP; 5394 A; 4676 C; 4890 G; 3546 T; 0 other;

Alignment Scores:

Pred. No.:	1,366-19	Length:	18506
Score:	565.50	Matches:	412
Percent Similarity:	36.99%	Conservative:	282
Best Local Similarity:	21.96%	Mismatches:	737
Query Match:	7.00%	Indels:	448
		Gaps:	81

US-09-902-432-4 (1-1596) x ABL05592 (1-18506)

QY 6 SerThngluGluInArgSerProGluInProAlaGlySer----- 18
 Db 10770 TCACAGAGAGCTCCCGGCGAGATCTGAGCGAGAAAGCTTGAAAGATCGCGAGACCG 10829
 QY 19 -----AspThrProSerGluLeuValLeuSerGlyHisGlyPro 31
 Db 10830 TCGCAGAGAGCTCTGCTCCGATCTGTACGAAAGCGTTAAGATGAAATCCCA 10889
 QY 32 AlaAla-----GluAlaSerGlyAlaAlaGly-----AspProAla 43
 Db 10890 GTCCGATCCCAAGAGAGCTTCAGACGATCCGTTGCGAAGAAATCCCAAGATTCGGCT 10949
 QY 44 AspAlaAspProAlaThrLys-----LeuProGlu---LysAsnGlyGluLeuSer 59
 Db 10950 GATGATTCAAAGAGACACGACGAGATCGTTCCTCAAGCGAAGCGGAGATATTAA 11009
 QY 60 SerValAsnGlyValAlaGluGluGluAspValHisValGluGluGluAsnGluGly 79
 Db 11010 GAT-----GAAATATCACTTTGGCTTAAAGATGAGCTGAAAGTCT 11054
 QY 80 GlnGluGluGlu-----ValValAspGluAspVal 89
 Db 11055 AAGGAGAAATCAAGCGAGATCGTGCGCGAACAATTCCTTGTCTCCAGAGAAATT 11114
 QY 90 GlyGlnArgGluSer-----GluAspValArgGluLysAspArgValGlu 104
 Db 11115 TCCAGACCGAGATCCGTTGCGAAGAGTGTAAAGACGAAAGCTGATAAGAA 11174
 QY 105 -----GluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAsp 119
 Db 11175 TCACCTTTAATGTCTTAAGAGAGCTTCACAGACGATCCGTTGCGAAGATGTTAAGAA 11234
 QY 120 GlyGlnGluGluThrSerGlu-----IleIleGluGluIleProAla 133
 Db 11235 GAACTGAGAGAGTCCAAAGAGAAATCTAGGCGAGAGTCACTGCTGAAAATACCTTTG 11294
 QY 134 SerGluAsnAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGly 153
 Db 11295 CCTTCCAAAGAGAGCTTCAGACGAGCTCCGTTGCGAAGAAAT----- 11336
 QY 154 PheLysLysValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGlu 173
 Db 11337 -----GTAAGAGATGAAAGCTGAT-----GluGly 190
 QY 174 LysSerAspThrValGlnLeuLeuThrValLysLysAspGluGly-----GluGly 190
 Db 11355 AAGTCC-----AAGAGAGAAATCTAGGCGAGAGTCCGGG 11387

QY 191 AlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGly 210
 Db 11388 GCGGAAAATCT-----CCTTTGGCTCCAGAAAGCTTCC 11423
 QY 211 GluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLeu 230
 Db 11424 AGACGAGCTTCCTGCGAAGAAATTAAGACGAGCTGAAAGTCTAAGAGAAATCT 11483
 QY 231 LysGlnGluGlnSerSerThrGluIleProLeu-----GlnAlaGluSerAspGln 247
 Db 11484 AGCGAGAGTGGTGTGCGAAGAAATCTCTTGGCTCCAGAAAGCTTCAGACCAACT 11543
 QY 248 AlaAlaGluGluGluAlaLysAspGluGlyGluLysGlnGluLysGluProThrLys 267
 Db 11544 TCCGTTCGAGAAAGTGTAAAGATGAACT--GAGAAAGTCTAAGAGAAATCTAAGAA 11600
 QY 268 SerProGluSerProSerSerProValAsnSerGluThrThrSerSerPheLysPhe 287
 Db 11601 GACTCGGTGGCGGAAATCTCTTGGCTCCAGAAAGCTTCC----- 11645
 QY 288 PheThrHisGlyTyrPalaGlyTyrPalaGlyLysThrSerPheLysSerLysGluAsp 307
 Db 11646 -----AGACCAGCTTCGTTGCGAAGAGTTCAGAGAC 11678
 QY 308 AspLeuGluThrAla-----GluLysArgLysGluGlnGluAlaGluLysValAspGlu 325
 Db 11679 GAAGCTGAGAGAGTCTAGGAGAAATCTAGCGAGAGTGGCGGCGAAGAAATCTCTTGT 11738
 QY 326 GluGluLysGluLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAsp 345
 Db 11739 GCTTCAGAGAAAGCTTCAGACGAGCTTCGTTGCGAAGAAATTAAGACGAGCTGAG 11798
 QY 346 GlnAlaArg-----LeuSer 350
 Db 11799 AAGCTTAAGAGAAATCTAGCGAGAGTGGGCGGCGAAGAAATCTCTTGGCTCCAA 11858
 QY 351 AlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAsp---LeuGluAla 369
 Db 11859 GGAGCTTCGAG--ACCCAACTTCGTTGCA--GAAAGTCTAAGAGAAAGCTAAGAG 11914
 QY 370 SerSerGluGluLysCys-----AlaProLeuAlaThrGluVal 382
 Db 11915 TCTAAGAGAAATCTAGTCGAGACTCGGTGCGGCGAAGAAATCTCTTGGCTGCTC 11968
 QY 383 PheAspGluLysMetGluAlaHisGlnGluValAlaAlaGluValHisValSerThrVal 402
 Db 11969 -----AAGGAGCTTCAGACCAAGCTTCGTTGCGAAGAGTTCAGAGCGAAGCT 12019
 QY 403 GlyLysThrGluGluGluGluGluGlyGlyGlyGluAlaGluGlyGlyValValGlu 422
 Db 12020 GAGAGTCTAAGAGAAATCTAAGCGAGAGTGGCGGCGAAGAAATCTCTTGT----- 12073
 QY 423 GlyThrGlyGluSerLeuProProGluLysLeuAlaGluProGlnGluValProGlnGlu 442
 Db 12074 GCTTCAGAGAGCTTCAGACCAAGCTTCGTTGCGAAG-----AGCTTAAGATGAA 12127
 QY 443 AlaGluProAlaGluGluLeuMetLysSerArgGluMetCysValSerGlyLysPheHis 462
 Db 12128 GCTGAGAGTCCAAAGAA-----GATCTAAGCGAGAGTGGTGGCC----- 12169
 QY 463 ThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHis---ProGluGly 481
 Db 12170 ---GAAATATCTCTTGGCTTCAGAGAGTTCAGACGCTTCGTTGCGAAGAAAT 12226
 QY 482 IleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGly----- 499
 Db 12227 GTTAAGATGAAAGTCTAAGAGAAATCTAAGCG--AGAGTCTGGTGGCGGAGAA 12285
 QY 500 -----SerProLeuLysLysLeuPheSerSerSerGlyLeuLysLysSerGly 516
 Db 12286 AATCTCTTGGCTTCAGAGAGCTTCAGACCAAGCTTCGTTGCGAAGAAAGTGTAAAG 12345

Qy	517	LysLysGlnLysGlyLysArgGlyGlyGlyAspGluGluProGlyGluThrGlnHis	536
Db	12346	ATGAAGCTCAGAAGTCTTAAGGAAGAATCAAGGCGA-GAGTCGGTGGCGCAAAATCTTCT	12404
Qy	537	IleHisThrGluSer-----ProGluSerAlaAspGluGlnLysGlyGluSerSer	553
Db	12405	TTGGCCTCCAAGAAAGCTTCCAGACAGCTTCCGTTGCAGAAAGTGTAAAGATGAAGCT	12464
Qy	554	AlaSerSerProGluGluProGluGluThrCysLeuGluLysGlyGlyProLeu-----	571
Db	12465	GAGAAGTCTAAGGAAGATCTAGGCGAGTCCGTTGGCGCGCAAAATCTCTTTGGCCTCC	12524
Qy	572	--GluAlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLys	590
Db	12525	AAGGAAGCTTCCAGACCAAGCTTCCGTTGCAGAAAGTGTAAAGCAAGCTGAGAAGTCT	12584
Qy	591	ArgGluGlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgVal	610
Db	12595	AAGGAAGAAATCTAGGCGAGAGTCGGTGGCGCAAAATCACCTTTGCCCTCCAAAGAAGCT	12644
Qy	611	ArgArgProSer-----GluSerAspLysGluGluGlu-----	621
Db	12645	TCGAGACCAACTCCGTTGCAGAAAGTCTAAGGATGAGCTGATAGTCCAAAGAGAA	12704
Qy	622	-----LeuGluLysValLysSerSerAlaThrLeu-----SerSerThr	633
Db	12705	TCTAGGCGAGAGTCGGGGCGCAAAATCTCTTTGGCTCCATAGAGCTTCCAGACCA	12764
Qy	634	AspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGlu-----	650
Db	12765	ACTTCGTTGCAGAAAGTGTTAAGAGCAAACTCAGAGAAGTCTAAGGAAGAATCTAGCGA	12824
Qy	650	-----	650
Db	12825	GAGTCGGTGCAGCAAAATCTCTTTGGCTCCAGAGGAGCTTCCAGACCAACTTCGTT	12884
Qy	651	-----GlnLysProGluProLysArgArgValAspThr	662
Db	12885	GCAGAAAGTGTTAAGGATGAAGCTCAGAAGTCTAAGGAAGAATCTAGGCGA-----	12938
Qy	663	SerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArgAlaArgLysAla	682
Db	12939	TCGGTGGCGCAAAATCTCTTTGGCTCCAGGAATCTCCAGACCAGCTTCGCTTGCA	12998
Qy	683	SerSerAspAspGlu-----	688
Db	12999	GAAAGTATTAGGACGAAGCTGAGGGGACTAAGCAGGAACTTAGCGGAGAGTCGATGCC	13058
Qy	689	-----GlyGlyProArgThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLys	706
Db	13059	GAAAGCGTGAAGCGGAGAGATCAAGGGTCA-----CAATCTCTTTAGCGTCCAAA	13112
Qy	707	AspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGln	726
Db	13113	GAA---ACGTCGAAGCCGATCTGTTGTGAAAGCGTAAAGGACGAACATGAAAGCCA	13169
Qy	727	GlySerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrp	746
Db	13170	GAAGGTCAGCATCCACAAAGTCGAGGTGCATCCCGACAGAGTCCGTGGCAGTA---	13226
Qy	747	GluSerPheLysArgLeuValThrProArgLysLysSerLysLysLeuGluGluLys	766
Db	13227	--AGTGTGAAGGACGAAGTCAACCGCTTCATTTCGGTCTCTGAATCAGTTGCCGACA	13283
Qy	767	AlaGluAspSerSer-----ValGluGlnLeuSerThr	777
Db	13284	TCGCCAGATGCTCGAAGGAAGCATCTCGTTCTTTATCGGTTGCCGAACCGCCTCATCA	13343
Qy	778	GluIleGluProSer--ArgGluGluSerTrpValSerIleLysLysPheIleProGly	796
Db	13344	CCCATTTGAAGAGTCCACGATCATCTGCTGACCTCAGTCTGCCCCCTTAACCTTACAGG	13403
Qy	797	ArgArgLysLysArg-----AlaAspGlyLysGln	806

	13403		13404	GAAGCTAAGGGCAACGCTTCCAACCTCTTTTCGAGTGCCATTGTATGCGCAAGAAGGACTTTT	13463
			:::	:::	
Db		807	GluGlnAlaThrValGluAsp-	-SerGlyProValGluIle	819
		:			
Db		13464	TTAGAAGTAAGCGGAATCCTCCCGCGGCCAGCTGTACTTTCAAACCACCGGAGTTT	13523	
		:			
Qy		820	AsnGluAspAspProAsnValProAlaValProLeuSerGluTyAsnAlaVal---	838	
		:			
Db		13524	TGCAGCCGATACTAGGCGCATACAGCTCCACGCCGTCGACGAGCTTCTCCCGTTCTG	13593	
		:			
Qy		839	--GluArgGluLysMetGlu-----	AlaGlnGlyAsnThr	849
Db		13584	CAGGAATCGAAGTAGTTGAACACAGCATACCACCTTCTGAGTCGGGCTACGGCGCAACA	13643	
		:			
Qy		850	GluLeuProGlnLeuLeuGly-----	AlaValTyValSerGlucIleu	864
Db		13644	CGAGAACCGGACCTTATGACCTGACTGAACCAAAATCCGACAGTGTAAACCAAGCAATCG	13703	
		:			
Qy		865	SerLysThrLeuValHisThrValSer-----	ValalaValileAsp	878
Db		13704	GAGACAACGTTATTTGAACCCCTTACCTCGAAGGTTGAAGAGTGAAGTACTGGAG	13763	
		:			
Qy		879	GlyThrArgAlaValThrServValGluArg-----	SerProSerTrpIleSer	895
Db		13764	AGC-----TCGGTGAAACAAGTTGAAGAGAAAGTCCAGACCTCAGTAAAGCAAGCAGAA	13817	
		:			
Qy		896	AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGlu	915	
Db		13818	ACCACTGTGACCGGATCTCTCGAAACAACCTAACCAGAGAGCAGTGACAGCTGTGACAG	13877	
		:			
Qy		916	Val-----ThrGluLysAspIleAlaGluGluThrPrValLeuThr	930	
Db		13878	ATCAAAATCAGTGTAGACACCAACTTTGAGAGAGTAGGCAAAAATCGTTGCCGATGTGCC	13937	
		:			
Qy		931	GlnThrLeuProGluGlyLysAspAlaHisAsp-----	AspMetValThr	945
Db		13938	AAGGTACTTAAGTCTGACAAGGACATTACAGACATCATACCCGACTTTTGCAGCAACGGCAA	13997	
		:			
Qy		946	SerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg	965	
Db		13998	CTGGAGGAGAACTGAAAGTCAACTGCCGATACAGAGGAGAGTCCGACAAAGTAGTACCCGG	14057	
		:			
Qy		966	ThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaVal	985	
Db		14058	GACGAGAAATCCCTGGAATCAGCGTCAAGTAGNA-----	ATC	14096
		:			
Qy		986	SerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValGlu	1005	
Db		14097	GAATCTGAGAAGTCATCGCCCGAT-----CAGAAGTCGGGCCCATTTCCATCGAGGAG	14150	
		:			
Qy		1006	SerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeu-----	1022	
Db		14151	AAGCACAAAAATCGAGCAATCAGAAAAACACATGTTGAGCGAGGAATAATTGACCAGCAGT	14210	
		:			
Qy		1023	----GlnAlaValAla-----AspLysValLysGluGluSerGlnValProAla	1037	
Db		14211	CGGCGGAGTCAGTTTCCAGCGCCGAGTGGTGCCCTCCCTTCCAGTCACGACGCC	14270	
		:			
Qy		1038	ThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu	1057	
Db		14271	AGCCATGAACATAAGGAGGTGAGTATCCGAGAGCCATAAAGCTGAGAAGTCTTCTCGC	14330	
		:			
Qy		1058	AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln	1077	
Db		14331	CCCAGTCCGTTGCCAGCCAGGTGACGCGAAAAGGACATGAAGACATCGCGCCCCAGCATCC	14390	
		:			
Qy		1078	GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer	1097	
Db		14391	AGCACGCCAATTTAGCACCAAGGAGCGATGAG-----GAGACCACTGAGTCG	14441	
		:			
Qy		1098	LeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLys	1117	

Db 14442 CTA-----CTTCACTGCTTACTACCAAGACAGTGGAG 14477
 QY 1118 LeuGlnLeuMetGlnGlnAlaValAlaProGluSerSerGluThrLeuThrAspSer 1137
 Db 14478 ACGAAACAGATGAGAGAAATCGAGCTTGTGATGCTCTCCACTCACTCACTAAGTCC 14537
 QY 1138 -----GluThrAsnGlySerThrProLeu 1145
 Db 14538 ACAGTCTCTCCAGTCACTGTCATCTTCAAGTCCAGAGAGTCCAGCCAGCAATCTTTG 14597
 QY 1146 AlaAspSerAspThrLeuAspGlyThrGlnGlnAspGlyThrLeuAspSerGlnAspSer 1165
 Db 14598 AGTAGCAGTCTCAGGTGAGACACACTCCAGCGAGTCTTCCAGCTTGTGGG 14657
 QY 1166 LysAlaThrAlaAlaValAlaArgGlnSerGlnValThrGluGluGluAlaAlaThrAla--- 1184
 Db 14658 GAGAAAGAGAGATGCTTCCAACTAGTTGAGAGAGACACTAGCCCTGATCTC 14717
 QY 1185 -----GlnLysGluGluProSerThrLeuProAsnAsnValPro 1197
 Db 14718 CAGCTGAGAGAGCTGCTTGTGAGTCCAGGAAATGCTCTCAAGTCAATCTGTACCGAG 14777
 QY 1198 AlaGlnGluGluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGlu 1217
 Db 14778 ATTCACACACAGATCGCTCAGAAATCAATTAAGAAATCAAGATCTAGAGAAACCAA 14837
 QY 1218 LeuThrAla-----AlaAlaValProValLeuAlaLysThrGluValGlyGlnGlu 1234
 Db 14838 GTAACACGACGAGTTTACCACTACACACAGAGCGCCACCAAGACATAGCTTAAGAAA 14897
 QY 1235 GlyGluValAspThrLeuAspGlyGlyValLysGluGluGlnGluValPheValHis 1254
 Db 14898 ACTGTGCGCAGATTGTGTGCGACCGAGAAATGTTGTCGCCCAAGAGAGCTTC----- 14951
 QY 1255 SerGlyProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyVal 1274
 Db 14952 ---AGACATGAGCCACTAAGTGGCCGAC----- 14978
 QY 1275 AlaGlyCysGlnGluLys---GluSerThrGluValGlnSerLeuSerLeuGluGly 1293
 Db 14979 ---GACTGTGTTGAAGAAAGACACCGCAGCGCGCTCCAGACATCCGCTCCAGCGA 15035
 QY 1294 GluMetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGlu 1313
 Db 15036 GCGCTATTCTGTAGCAGCGAGATCACGGAGAGAGTCTGCTCAGTCAAGGAGATGAG 15095
 QY 1314 -----GluGlyGluGlnGluThrAlaAlaProGluHisGlu-GlyThrTyrGlu 1329
 Db 15096 AGCCGCTTGAACCAAGTATCTCTGAAGATGAGAAACCGCTATGATGATCGACGACGG 15155
 QY 1329 LysProValLeuThrLeuAspMetPro----- 1338
 Db 15156 T---CCAGTGTTAAGAGAGCGCTTCCAAAGTATTGCTACATCATGATGACAGCATC 15212
 QY 1339 -----SerSerGluArg-GlyLysAlaAla 1346
 Db 15213 TACAAACCTTCCAGAGATATGAAACGATTTCAAAGCTCTCTGAGAGAGACAGCAC 15272
 QY 1346 euGlySerLeuGlyLysProSerLeuProAspGlnAspLysAlaGlyCysIleGluVal 1366
 Db 15273 GTGGAGAGCTGGCTCAGAGAGTCACTTCAACAGCAAGACAC---AACTCTTTGGCAA 15329
 QY 1366 AlaGln-ValGlnSerLeuAspThrThrValThrGlnThr-----AlaGluAlaVal 1382
 Db 15330 TCCAGGCAAGACAGACAGCAACACACAGTAGCAGCAAGACCGGTGCTCAAGGTTG 15389
 QY 1383 GluLysValIleGluThrValValIleSerGluThrGlyLysSer----- 1397
 Db 15390 GAGTCCATTACCTTCAACAGATGATCAAGAACTAGCAGAGTCAAGGCGACCTTGGC 15449
 QY 1398 -----ProGluCysValGlyAlaHisLeuLeuProAlaGlyLysSer---Ser 1412
 Db 15450 GATCGCAAGACACCGGCACT-----GCGCAGATAGTCCCGGCGTCAAGAGCAATGAGC 15503

QY 1413 AlaThrGlyGlyHisThrThrLeuGlnHisAla---GluAspThrValProLeuGlyPro 1431
 Db 15504 TCGACAGATACACCGGCTCTGTGATTTGAGACTGAGCTGAGACCGCTGGTGA 15563
 QY 1432 GluSerGlnAlaGluSerIleProIleLeuValThrProAlaProGluSerThrLeuHis 1451
 Db 15564 AAGTGTAGTCAAGTCCCGCCAGCATCTGATCTTCAATCGGGCCGATGTCGCC- 15617
 QY 1452 ProAspLeuGlnGlyLysLeuSer-----AlaSerGlnArgGlu----- 1464
 Db 15618 AAGCATCATCAGCGCAATTAAGTCCCGCTGCTCAGCTGAGAGACAGCAATTCCTCC 15677
 QY 1465 -----ArgSerGluGluLysProAspAlaGlyProAspAlaAspGly 1480
 Db 15678 ACTCGGTAGACGGGAGTTCGATACGACACCCCGAATCTCGCCA----- 15725
 QY 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGlu 1500
 Db 15726 AAGCCCATCATCCTTTTCCCGCGCTCAGACAGACAGCTTAAGTCTTGAAATGCA 15785
 QY 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
 Db 15786 GTATGTTAAGCAGTTTGTGTGCTGATTTATTTGCTATTATATCTATGTTTAAAC 15845
 QY 1521 ---ThrGluThrAlaProGluThrHisAlaTyrAspSerGlnThrGluValPro----- 1537
 Db 15846 CCATGTAGACCATCTCGACGAGATGTTGGCAGGACGAGCTGTCAGCCGAGCTG 15905
 QY 1538 AlaCysArgLeu-AspSerArgGluProAsnArgCysThrThrLysMetLysAspAlaLys 1557
 Db 15906 AGTGTAGGAGTCACTTCCAGTCTGACAGAACTCCGTGACTAAGTGT----- 15954
 QY 1557 MetLysHisProValProGlnProArgGluAspLeuGln 1570
 Db 15955 -----CACACAGCCCTCAG---CGGAGACACGACAA 15984
 RESULT 12
 ABL19989
 ID ABL19989 standard; DNA; 20448 BP.
 AC ABL19989;
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 11440.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ds.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 PA *PEPE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI; 2001-656860/75.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 11440; 21bp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS7202).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 20448 BP; 7528 A; 4083 C; 4878 G; 3959 T; 0 other;

Alignment Scores:

Pred. No.: 3,9e-19 Length: 20448
 Score: 557.50 Matches: 396
 Percent Similarity: 36.48% Conservative: 300
 Best Local Similarity: 20.75% Mismatches: 725
 Query Match: 6.91% Indels: 487
 DB: 23 Gaps: 79

US-09-902-432-4 (1-1596) x ABL19989 (1-20448)

QY 7 ThrGluGlnArgSerProGluGlnProAlaGly----- 17
 DB 8965 ACTACAAAGCAATGTTGGACACCTGCTGGCGAAGGTGAATAGAGATTATTGAAGTT 9024
 QY 18 -----SerAspThrPro-----SerGluLeuValLeuSerGlyHisGlyProAlaAlaGlu 34
 DB 9025 GTTAGGGATGATCAGCTTGGCGAGCAAAATACTATTGTGGATGATGAGCA----- 9075
 QY 35 AlaSerGlyAlaAlaGlyAspProAlaAspAlaAspProAlaThrLysLeuProGlnLys 54
 DB 9076 -----GAACTGTGAACCAAGACAGCAAAACCCCAAGAACTTAAGAAG 9117
 QY 55 AsnGlyGlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHis-----Val 72
 DB 9118 AAACCCAGG-----AAGGTTAAAGAGATGATATACATGACTACATA 9159
 QY 73 Gln-----GluGluAsnGlnGluGlyGln 80
 DB 9160 CAAAGTTAATTGAGTTGGAAACCAACCAAGACCGAGCTTGAAGATATGAGAAATTGAA 9219
 QY 81 GluGluGluValValAspGlu-----AspValGlyGlnArgGlu 93
 DB 9220 TTGCAACCAATAGTCAAGGACACAGCCATCGATAGCCCGATTCACGTTTGGACGAATCG 9279
 QY 94 SerGluAspValArgLysAspArg----- 102
 DB 9280 CCAAAAGAGTGCAGAAAAGATATAAAAGTCAAGGTCTACAAAAGTTCGAATGAGGAA 9339
 QY 103 -----ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
 DB 9340 ACTCTGTTTCAGAACCAATATGCAAGGTCAATGTTGTGCAAGAACAGCTCTGAGCAA 9399
 QY 121 GlnGluGluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGlu 140
 DB 9400 CCGGAATACCAAGTGCAGAAATCTAGAAGTTAAGCCCGTA---GAGGTAGATGTTAAAGAG 9456
 QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
 DB 9457 GTTATAACTGAAGATGCGAAGCCGGTTCAAGAAAAGACTACTAAGCGAGTTTAAAGAAG 9516
 QY 161 ValGly-----PheLysPheThrValLysLysAspLysAsnGluLys 174
 DB 9517 ATTGGCCCGGAGAACCAACCTTTAAGATTACAAATG-----ATTGAAGCCGAGGAC 9570
 QY 175 SerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGlyAlaGluAlaSer 194
 DB 9571 AAGCATTCGGTT-----ACAGTTATTGTTGACGAGAACCCGGAATAGCTTCTCCA 9621
 QY 195 ValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGluSerAlaSer 214

DB 9622 CAGTCTATCGAAGAACATCCGAAACAGTCTTAAGGAAAGAGTCTCAAGGCCCAAGAAA 9681
 QY 215 -----LysGluSerGluLeu----- 219
 DB 9682 ACTGTGCAGAAAAGTTAAAGAGCAGCTTAAGGATTACGTAAGAACTAATTTGAAGAA 9741
 QY 220 -----LysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSer 236
 DB 9742 GAAATTCCTTAAGTTGACCTTGAGAAATATGAA-----AAGGTAGAAATGCCAGAA 9792
 QY 237 ThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGlu 256
 DB 9793 AAACCGATTAAACTACTGTT---TCTGATTCCATTCGGAAGAACCAAACTGACAAA 9849
 QY 257 GlyGlu-----GluLysGlnGluLysGluProThrLysSerProGlu 270
 DB 9850 TCCCAACCAATATCAGTACTACCCGACACACGAAACCCCAAGAAACGAGACTCCCAAG 9909
 QY 271 SerProSerSerProValAsnSerGluThrThrSerSerPheLysLysPheThrHis 290
 DB 9910 ACTCCAAAGACAGACACAGACACAGACCAAGTTCCGAGACGAAACCCGACAGACCACTGTA 9969
 QY 291 GlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspLeuGlu 310
 DB 9970 GACACTACTGACATACAGAGCTTACCAACCCCAACTGCACAGCCCGGAGGACACAGCC 10029
 QY 311 ThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluGlu----- 327
 DB 10030 ACTGCACAATTTACCAACGCGCACCAAGAGAGAAATCTTACTCAAGTACACTAAAGAT 10089
 QY 328 -----LysGluLysThrGluProAlaSerGluGluGlnGluPro 340
 DB 10090 ACAATTCGAAACACAGTTAAACATAAGAAAACAAACCA-----GACACACAAAAAGC 10143
 QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr----- 353
 DB 10144 GTTGAACCAAGTGAAGTACCAGAGGTTTCAATGAAGACTATCAAAATTAGCATCATCCAGAA 10203
 QY 354 -----GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
 DB 10204 GAGCTTGTGCGAGAGAGAGAGCCGCGGAAAGATTTTGGAGGTTAGAGTTATCGAT---GAA 10260
 QY 369 AlaSerSerGluLysCysAlaProLeuAlaThrGluValPheAspGlu----- 385
 DB 10261 GTCGCGGAGGTGGAAGAATCACAGCCCATTTGTGGAAGAGGTAGAAGATCAGGAACACAG 10320
 QY 385 ----- 385
 DB 10321 CCAGCTACAGAAAGAACTGTTGAGGACGCTCACTAAGCTTAAAGTCTAAAAAGAGAGGTG 10380
 QY 386 -----LysMetGluAlaHisGlnGluValValAlaGluValHisValSerThrVal 402
 DB 10381 GTTAAAGAACAGACCGATGACCATGACGACTCATCAAAAAGATGTTGGAGCAGGAAATA 10440
 QY 403 GluLysThrGluGluGlnGlyGlyGlyGlyGluAlaGluGlyGlyValValValGlu 422
 DB 10441 GAGAAGACGAACTCGAAAAATACGAGAAAATTGAATTTGATGTTCCCAAGAAACTCAAA 10500
 QY 423 GlyThrGlyLysSerLeuProGluLysLeuAlaGluProGlnGluValPro----- 440
 DB 10501 CCAGAAATTCGCTGCTCTCGAACCAATAAAAAATAGAGCGCAAGAACAGAACCAAG 10560
 QY 441 -----Gln 441
 DB 10561 GTGACTATTCTCGATGCCACTGATGTTCTTAAACCGTGAAACTAAACCGTAGTAACGT 10620
 QY 442 GluAlaGluProAlaGluGluLeu-----MetLysSerArg 453
 DB 10621 AAGGAGAAACCTGCGGAGAACTGACCGTTCACTTCCAAAGTTCCAGACTTAAGCCCGT 10680
 QY 454 GluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGluLys 473

QY 1096 GluSerLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnVal 1115
Db 12661 GTGTTCAAGTCAAGCTCTGAGAGTAAGATTGTAGAGGAAGAGTATTGAGAGGAAG 12720
QY 1116 IleLys-----LeuGlnGlnLeuMetGluGlnAlaValAlaProGlu-----SerSer 1131
Db 12721 CCTAAGAGTTCAATTCGGTTTCTGAATCAGAACCCAGCCGAGGAGCCAGTGTC 12780
QY 1132 GluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp----- 1147
Db 12781 GAGCAGTTACAGTAAATAAAGCAAGCCCTCTGTTACTTTCGCGAGCAACCTGCCACT 12840
QY 1148 -----SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAsp 1161
Db 12841 GAAATCGTTATCAGGAAGAACGCGCTGAAGTTGCTACTGAAGACGCTCATATCAAG 12900
QY 1162 SerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluAla 1181
Db 12901 ACCAAGAGCCAAAG-----AAGAGGTAAACGGATGTGGAAGCT 12939
QY 1182 AlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGlu 1201
Db 12940 GAGGAGCTAAAGATTAAAGTCACTGAGGAAGTCCACAGGAATTCCTATATTGGAGGAG 12999
QY 1202 HisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluThrAlaAla 1221
Db 13000 GTCTCGGAGGAA-----GAGGTCACTGAAACCAAGAG-----ACA 13038
QY 1222 AlaValProValLeuAlaLysThr---GluValGlyGlnGluGlyGluValAspTrpLeu 1240
Db 13039 GCGCTCTGTTGAAGAGAGACATACAAATTTGGTATCAAGGAACCGGAGCCGGAAG 13098
QY 1241 AspGlyLysValLysGluGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
Db 13099 CTGCGCAAGCAATTGCGAGGAGGAGCGCGTGTACAGAGCCCAATTTGAGGAAGCA 13158
QY 1261 LysAlaAlaAspValThrTyArgSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
Db 13159 CCGAAACCGGAAGTCTTTGAGGAACACAAAGTTAGATCATT----- 13200
QY 1281 GluSerThrGluValGlnSerLeuGluGlyGluGlyGluMetGluThrAspValGlu 1300
Db 13201 -----GAGGAACACCTCGGGAATTGTCGAGGAAGTCATTGAGGAGGAGGTGAAG 13251
QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGlyGluGlnGlnThrAla 1320
Db 13252 GTAATTCGAGAGAAGAGCTTAAGCAGAAATTAAAGAGGAACCCGAGCGCGAGTTACT 13311
QY 1321 AlaProGluHisGluGlyThrTyArgProValLeuThrLeuAspMetProSerSer 1340
Db 13312 GTT-----TCTACTCCAAAGCCTGTC----- 13332
QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLys 1360
Db 13333 GAAGAAGTGAAGTACGCTCCAGTATTGCTGTTATTCAGAACCAACCCGACTGAGGAGGAA 13392
QY 1361 AlacGlyCysIleGluValGlnValGlnSerLeuAspThrValThrGlnThrAlaGlu 1380
Db 13393 GCGCTGATCTCAGATAACATCATTTGAGAAGAA-----ACACCACCCACAGAA 13443
QY 1381 AlaValGluLysValIleGluThrValValIleSerGlu----- 1393
Db 13444 CTAGTTCCAGGAATTGAAGAAATTGAATTCGTTGAGGAACCAAGGCCCTGGAAGAACAA 13503
QY 1394 -----ThrGlyGluSerProGluCysValGlyAlaHisLeuLeu 1406
Db 13504 CCAACGGATTTCACCTTCGCCACAAAGGATTCGAAAGAACCGACTGTGGAAGAGCTG 13563
QY 1407 ProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThr 1426
Db 13564 CCGAGGACACAGTGC-----ACGATTCAAAAAAAGAAAGAG 13602
QY 1427 ValProLeuGly-----ProGluSerGlnAlaGluSerIleProIleIleValThr 1443

Db 13603 GCACCAAGTACCAAGAGTAGTTGAAGAACCCGAGCTGAGTTTGTGTTAAGCCAAAGACG 13662
QY 1444 ProAlaProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArg 1463
Db 13663 CCTGTCCAAGAGTTACT-----GAAGAAGCTAAGATAACGAAGTCTAAGAAA 13710
QY 1464 GluArgSerGluGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSer 1483
Db 13711 CCTGTAAGAGGAAGAG-----GCCGCGGAGAGCTTAAA 13746
QY 1484 ThrAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeu----- 1499
Db 13747 GTAACGATCACTGAAGAAATTCCTCACTGAACCAAGATTCAAGAAATTTATCGAGGAGATC 13806
QY 1500 -----GluSerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaVal 1515
Db 13807 GAAGAGATCGAGGAAGAAAACCCGAGAGTAGTGTGATCGAAGTTAAGGAAACCAACCA 13866
QY 1516 AspGlnPheAlaArgThrGluThr-----AlaProGluThr 1527
Db 13867 GAAGCTGTAGAGATAAGGAAGTTAGCTTACCAGAAAAGAAACCTAAGGCTCCAATAGTT 13926
QY 1528 HisAlaTyArgSerGln---ThrGlnValProAlaCysArgLeuAspSerArgGluPro 1546
Db 13927 GAAGACCCAGAGCTGAAATAACTTTGAAGCCAAAGGTAAAGTCGGAAGAAAGTTCAGGAA 13986
QY 1547 AsnArgCysTrpThrLysMetLysAspAlaLysMet----- 1558
Db 13987 GAAGCAAGATAGTAGAAGAAAAGCAAGAAAATCGACGAAGTTGCAGTAGCCGACGAG 14046
QY 1559 -----LysHisProValProGlnProArgGluAspLeuGlnValLeu 1572
Db 14047 CTTACGCTCAAGGTTCAAGAGAGAGGTTGTCCTCAGAGCCGATTGTTGAAGAAGAAGTGATA 14106
QY 1573 ThrValLeuGluAlaTrpAlaGlnProArgLysCysLeuProArg----- 1587
Db 14107 GAAGAGTTCCAAATTAAAAAGAAACCAAGAACCCGAGCCAGAGGACATTTGTCGATGCT 14166
QY 1588 -----LeuGlnLeuLysAlaPro 1593
Db 14167 GCGATTGTGAATCAAGAAACCA 14190
RESULT 13
ABL19988/c
ID ABL19988 standard; DNA; 24971 BP.
XX ABL19988;
AC ABL19988;
DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 11437.
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 11437.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX WO200171042-A2.
PN 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX

QY 544 SerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluProGluGluThr 563
 Db 10534 -----GAAAGAAAGAGTCGGAAGAAACCAAAATCTTACGAGTTCAAAATTTACAGAGCG 10481
 QY 564 ThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGluGlyThr 583
 Db 10480 CAATCTATTGAAGAAAGCCCAATAGAGTGGCTGAAGAGCTCTCGAGGA----- 10430
 QY 584 ThrSerAspGlyGluLysLysArgGluGlyThrProTrpAlaSerPheLysMet 603
 Db 10430 ----- 10430
 QY 604 ValThrProLysLysArgValArgProSerGluSerAspLysGluGluGluGlu 623
 Db 10429 ---ACACCAAGGTGTAGAGAAAGGTTGCCAGAAATTCGATTCCTACGAATTC--- 10376
 QY 624 LysValLysSerAlaThrLeuSerSerThrAspSer-----ThrValSerGlu 639
 Db 10375 -----ACTCTAAAGAACTGATGAAGAAAGGTAATCACAGTTGACGAC 10331
 QY 640 MetGlnAspGluValLysThrValGlyGlu-----GluGlnLysProGluGluProLys 657
 Db 10330 CAGCCCGAGGAAGAGCCCTGTAGAGTGTATTCAAGAAAGAAACCAAGAACCTTGAA 10271
 QY 658 ArgArgValAspThrSerValSerTrpGluAlaLeuLeuCysValGlySerSerLysLys 677
 Db 10270 GCAGTTGAGCCCGAGTTGTAAATGACTCAACCCCAAGATTGTGGAGAAACGAGCGTGAA 10211
 QY 678 ArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArgThrLeuGlyGlyAsp 697
 Db 10210 ACTGCAATCAAA----- 10199
 QY 698 SerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAla 717
 Db 10198 CAGAAAAAGCAGAAAAACCGAAGAGCAGAGAGAGTCAATTAGCAATCAAGTT 10139
 QY 718 SerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerPro 737
 Db 10138 GTCGAAAGTGAAGCCCGGTGCGGAGAGAGTGTTCAGAGTCTCCTGAGAGTAAGATT 10079
 QY 738 SerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLys 757
 Db 10078 GTAGAGAGAGATTATGCA----- 10058
 QY 758 LysSerLysSerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThr 777
 Db 10057 -----GAAGAAAGCCTAAGAGTTCAATTCGC-----GTTTCT 10022
 QY 778 GluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheLeuProGlyArg 797
 Db 10021 GAATCAGAACCCAGCCCGAGGAGCCCA-----AGTGTGAGCAGATTACAGATTAAAAA 9968
 QY 798 ArgLysLysArgAlaAspGlyLysGlnGluAlaThr-----ValGluAspSer 814
 Db 9967 CGCAAGCCTTCTGTACTTTCGCGAGCAACCTGCCACTGAAATCGTTATCAAGGAAGC 9908
 QY 815 GlyProValGluIleAsnGluAspProAsnValProAlaValProLeuSerGlu 834
 Db 9907 AAGCCCGCTGAAGTTGTCAGTGAAGAGCTCATATCAAGACCAAGAGCCCAAGAGAG 9848
 QY 835 TyrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeu 854
 Db 9847 GTAACCGATGTGGAAGCTGAGGAGCTAAAGATTAAAGTCACTGAGGAAGTTCCACAG--- 9791
 QY 855 LeuGlyAlaValTyrValSerGluLeuSerLysThrLeuValHisThrValSerVal 874
 Db 9790 -----GAAATTCCTATTTGAGGAGGTTCTCGAGAA----- 9758
 QY 875 AlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIle 894
 Db 9757 GAGGTGATCACTGAAACCAAGAGACAGCGCTGTTGTTGAGAGAGACATACAAATTT 9698
 QY 895 SerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGlu 914

Db 9697 GGTATCAAGAAACCGAGCCGGAAG-----CTGCCGAA 9662
 QY 915 GluValThrGluLysAspIleAlaGluGluThrProValLeuThrGlnThrLeu--- 933
 Db 9661 GCAATGTTCGAG-----GAGGAGAGCCGGTGTACAGAGCCAAATTGAG 9617
 QY 934 -----ProGluGlyLysAspAlaHis----- 940
 Db 9616 GAAGCACCCGAAACCGAAGTCTTTGAGAAACACAAGGTGTAGAGTCATTGAGAAACACCT 9557
 QY 941 -----AspAspMetValThrSerGluValAspPheThrSerGluAlaValThr 956
 Db 9556 CGGAATTTGTCAGGAAGTCATTGAGAGAGGTGAAGTAATTCGACAGAAAGACCT 9497
 QY 957 AlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGlu 976
 Db 9496 AAGCCAGAAATTAAGAGGAAACCGAGCGGAGGTTACTGTCTTCTACTCCAAAGCCTGTC 9437
 QY 977 GluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGlu 996
 Db 9436 GAAGAAGTGAAGACTAGTCCAGTATTCCTGTTATTCCAGAAACACCG---ACTGAGGAG 9380
 QY 997 GluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluArg 1016
 Db 9379 GAAGCCGCTGATCTCAAG-----ATAACAATCATTTGAAGAAGAAACA 9338
 QY 1017 GlnThrGlnAlaIleLeuGlnAlaValAlaAsp---LysValLysGluGluSerGlnVal 1035
 Db 9337 CCACCACAGAACTAGTTCAGGAATTTGAAGAAATTTGAATCTGTTAGGAACCAAGGCC 9278
 QY 1036 ProAlaThrGlnThrValGlnArgThr---GlySerLysAlaLeuGluLys----- 1051
 Db 9277 CCTGAAGAACCAACCAACCGGATTTCACTTCGCCACAAAGGATTCGGAAGAAAGCCGACT 9218
 QY 1052 ValGluGluValGluGluAspSerGluValLeuAlaSerGluLysGluLysAsp----- 1069
 Db 9217 GTGGAAGAGCTGCCCGAGGAACAGGTCACGATTCAAAAAAGAAAAAGAAAGGCCACAGTA 9158
 QY 1070 -----ValMetProLysGlyProValGln 1077
 Db 9157 CCAGAGGTAGTTGAAGAACCCGAGCTGAGTTTGTGTTAAGCCAAAGACCCCTGTCCAA 9098
 QY 1078 -----GluAla 1079
 Db 9097 GAAGTTACTGAAGAGCTAAGATAACGAGTCTAAGAAACCTGTAAAGGAAGAGGCC 9038
 QY 1080 GlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGlu 1099
 Db 9037 CGGCGAGAGCTTAAAGTAAAGTCACTGAGAAATTTCCCACTGAACCAAGAGTTCAAGAA 8978
 QY 1100 ValProGluValThrAlaAspValAspHisValAlaThrCysGln---ValIleLysLeu 1118
 Db 8977 ATTATCGAGGAGATCGAAGAGATCGAGAGAAAGAAACCCCGCAGAGTATGTATCGAAGTT 8918
 QY 1119 GlnGluLeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGlu 1138
 Db 8917 AAGGAAGCCCAACAGAGCTGTA-----GAAGATAAGGAAGTGTAGTACCAGAA 8867
 QY 1139 ThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGlu 1158
 Db 8866 AAGAAACCTTAAGGCTCCAATAGTTGAAGAGCCAGAGCTGAAATAAATCTTTGAAGCCAAAG 8807
 QY 1159 ThrIleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGlu 1178
 Db 8806 GTAAAGTCGGAAGAGTTCAGGAAGAAAGCAAGATGATGAAGAAAGAAAGCCCAAGAAATC 8747
 QY 1179 GluGluAlaAlaThrAlaGln-----LysGluGluProSerThrLeuProAsn 1194
 Db 8746 GACGAAGTTGCATAGCCGAGAGCTTACGGTCAAGGTTGAAGAGAGGTTGTGCCAGAG 8687
 QY 1195 AsnValProAlaGlnGluHisGlyGlu---GluProGlyArgAspValLeuGluPro 1213


```

Db 8686 CCGATTGTTGAAGAAGATGTAGACAGTTCGAATTTAAAAAGAACCAAGACCG 8627
Qy 1214 ThrgingluLeuThraAlaAlaValProValLeu----- 1226
Db 8626 GAGCCAGAGACATTTGTGATCTCGATTTGTAAGACTTAAGAAACCAAGACCGATAGAT 8567
Qy 1227 -----AlaLeuThrgluValGln 1233
Db 8566 GCTGATGAAGTTGTGCTGAAGTTACCTTAAGCCAAAGCGTAAACGAAATTTACAGAG 8507
Qy 1234 GluGly-----GluValAspTrpLeuAspGlyGluValValGlyGlu----- 1248
Db 8506 GAAAGATTTCTGTGATGTTAAGCTGCCAAGAAAGAAAGATTACAGAAATTTCA 8447
Qy 1249 ---GingluValPheValHisSerGlyProAsnSerGlnValAlaAspValThrTyr 1267
Db 8446 GACCAACCTGTACAGCTCAAGAAAAAGAAAGCCGAGAAACCGGTGGAAGAGCGTGG 8387
Qy 1268 AspSerGluValMetGlyValAlaGlyCysGlnGluValSerGluValGlnSer 1287
Db 8386 GACGAGCTTAAGCTA-----CAGCAAACTGTTGTGAGGAAAGACCT 8345
Qy 1288 LeuSerLeuGluGluGlyGlu---MetGluThrAspValGluValGlyValSerGluThr 1306
Db 8344 GTGGAAATTTGAAGAGAGATCTATTGAGAAAGCATTCGCAAAAAGCCAAAG 8285
Qy 1307 LysProGluGlnValSerGluGluGlyGluGluGluThraAlaAlaProGluHisGluGly 1326
Db 8284 AAGCCTTCGAGCGCAGCGGTGAGATCTCAAGAAACA-----GAGTTCAAGTCTC 8234
Qy 1327 ThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluValGlyValAlaLeu 1346
Db 8233 TCCTTCAAAAAGCCC---CATACATTAATGAGGTTGAGAGAGCCGCCACAGTTCTG 8177
Qy 1347 GlySerLeuGlyGlySerProSerLeuProAspGlnAspGlyAlaGlyCysIleGluVal 1366
Db 8176 AAAAAACGACCGTCAAGCCCTACTCTGATGAAGCTGCAGCT----- 8132
Qy 1367 GluValGlnSerLeuAspThrThraValThrglnThraAlaGluValGlu-----Lys 1384
Db 8131 -----GACTTTCTATTAAAGCTGCAAGAAAGAAATACGAAGAGGCGAG 8087
Qy 1385 ValIleGluThraValAlaIleSerGlu-----ThrGly 1395
Db 8086 GATATCGAAGAGTTCGCTGATGTCAGCAAGAAAGCCAGACCTTTGCMAATTCACAGA 8027
Qy 1396 GluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGly 1415
Db 8026 GAGGATGAGAGGCTTACACAGTGAAGAAACTTAAGCCGTCGAAAGCA----- 7979
Qy 1416 GlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAla 1435
Db 7978 -----GTAACATTCGCCGAATATGCTGATGTT 7952
Qy 1436 GluSerIleProIleIleValThrProAlaProGluSerThrLeuHisProAspLeuGln 1455
Db 7951 GAGAACGTCACATTC-----CGTGCAGAAAGTACCAAGACCAAGAGATGTCAT 7901
Qy 1456 GlyGluIleSerAlaSerGlnArgGluAspSerGluGluGlu----- 1469
Db 7900 CAAGAAATTCACATTCCTCGACATCTGATGCCGAGAGAAATTTCCATGTCGAAAG 7841
Qy 1470 -----AspLysPro----- 1478
Db 7840 GTGAAGTTAAAGAGCTTACAAAGACCTTCTCGAAGCAGCGATGACGCAAGATC 7781
Qy 1479 -----AspGlyLysGluSerThrAlaIleGluLysValLeuLysAla 1492
Db 7780 AAGATCATTCAGAGACTTTCATGATGGTGAAGAACCATTAATCGAAGAAATTCGAGAGAT 7721
Qy 1493 GluProGluIleLeuGluLeuGluSerLysSerAsnLysIleValLeuAsnValIleGln 1512
Db 7720 GAGGATACCATCATGATGATGAGGAAACCAAGAAAGATCTTTGTGAGGAATTTGCCACA 7661

```

```

Qy 1513 ThrAlaValAspGlnPheAlaArgThrGluThraAlaProGluThrHisAlaTyrAspSer 1532
Db 7660 GACCAAGAGACTT---CAAGTTAAGCCAAAGAAAGACCCCAACCAAGATATTCGTTCA 7602
Qy 1533 Gln 1533
Db 7601 AGA 7599

RESULT 14
ABL28640
ID ABL28640 standard; DNA; 24789 BP.
XX
AC ABL28640;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SHQ ID NO 37393.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
EN W0200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li FMD, Myers EW,
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 37393; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 24789 BP; 7654 A; 5448 C; 6281 G; 5406 T; 0 other;

Alignment Scores:
Pred. No.: 1.17e-17 Length: 24789
Score: 530.50 Matches: 389
Percent Similarity: 36.85% Conservative: 309
Best Local Similarity: 20.54% Mismatches: 639
Query Match: 6.57% Indels: 559
DB: Gaps: 79

US-09-902-432-4 (1-1596) x ABL28640 (1-24789)
Qy 8 GluGlnArgSerProGluGlnProAlaGlySerAspThrProSerGluLeuValLeuSer 27
Db 9664 GACCAAGAGTCC---CAAGAACCAAGATATTCGAGTCGTAAGCAAGATATCTGAA 9720

```

Qy	28	GlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAspAlaAspPro	47
Db	9721	CGAAAGATTGAAGAACCAAGAAACCTCGAAGAAATGGACACTGAAGCTTAAATCATCAGAGAA	9780
Qy	48	AlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAsnGlyValAlaGlu---	66
Db	9781	GCTACAGTCTTGGATACCAAGTCTCGAAGAAAGAAAGAACTAGAGGCTAGTGCAGAAAG	9840
Qy	67	GlnGlyAspValHisValGlnGluAsnGlnGlnGlyGlnGlyGlnGluGluGluValValAsp	86
Db	9841	CAAGTGATCAGGATGTCGAAAGAAGTCCCGAGAAACCCAGAAAGTATCTGAGGTCGTAGCT	9900
Qy	87	GluAspValGlyGlnArgGluSerGluAspValArgGlu-----	99
Db	9901	GAGAGATATCTGAAGAAACAATTGAAGAGAGCAAGAAACCTGAAGTAAAGCAGACCTGAG	9960
Qy	100	-----LysAspArgValGluGlu	105
Db	9961	ATTAAATCAGAGAAAGCCACAGCCTTGGATTAAGCAAGTCTTGGAAAGAAAGAACTAGAG	10020
Qy	106	MetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu---GluThr	124
Db	10021	GCTAGTCACAAAAGCAATGTGATCAGGATGTCGAAAGAAAGTCCGAGAAACCCAGNAGTA	10080
Qy	125	SerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMetValGlnPro	144
Db	10081	TCTGAGGTCGTAGCTGAG---AAAATATCTGAAGAACCAATTTGAAGAGCCAAAGAAACCT	10137
Qy	144	-----LysAspArgValGluGlu	144
Db	10138	GAAGTAAAGGAAACCTGAGGTTAAATCAGAGAAAGCTACAGTCTTGGATTAAGCAAGTCTCTG	10197
Qy	145	-----AlaGluSerGlnAlaAsnAspValGly---	153
Db	10198	GAAGAAAGGAACTAGAGGCTAGTGCAACAAAGCAAGTGTATCAGGATGTGCAAAAGNAG	10257
Qy	154	PhelLysLys-----ValPheLysPheValGlyPheLysPheThrValLys-----	168
Db	10258	TTCCAGAAAGCAGAAAGTATCTGAGGTCGTGAGGAGATATCTGAAAGAAATCTGAAAGAAACAATTCGAA	10317
Qy	169	-----LysAsp-----LysAsnGluLysSerAspThrVal	178
Db	10318	GAGCCAAAGAAACCTGAAGTAAAGGACACTCAGATTAATCAGAGAAAGCCACACGCTTG	10377
Qy	179	GlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGly	198
Db	10378	GATAGCAAGTCTCGAAGAAAGGAACCTAGAGGCTAGTGCAAAAGCAAGGCTGATCAG	10437
Qy	199	Asp-----HisGlnGluProSerValGluThrAlaValGlyGluSerAlaSer	214
Db	10438	GATGCGAAAAGAAGTCCCGAAGAACCAAGTATCTGAAGTCATAGCTGAGGAGATATCT	10497
Qy	215	LysGluSer---GluLeuLysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGlu	233
Db	10498	GAGGAAAGATGTTGAAGAACCAAGAAACCTCAAGAAAGAAAGAACTGAGGTTAAATCATCAGAG	10557
Qy	234	GlnSerSerThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAla	253
Db	10558	AAA---GCTACAGTCTTGGATTAAGCAAGTCTCGAAGAAAGAAAGAACTAGAGGCTAGTGCA	10614
Qy	254	LysAspGluGlyGlu-----GluLysGlnGluLysGluProThrLysSerPro---	269
Db	10615	CAAAAGCAAGGTGATCAGGATGTCGAGAAAGGTCCTCCAGAAACCCAGAAAGTATCTCAGGTC	10674
Qy	270	-----GluSerProSerSerProValAsnSerGlu	279
Db	10675	GTAGCTGAGAAGGTATCTGAAGGAAAGATTTGAAGAACCAAGAAACCTGAAAGTAAAGGAA	10734
Qy	280	ThrThrSerSerPheLysLysPhePheThrHisGlyTTPAlaGly---TrpArgLysLys	298
Db	10735	ACAGAGCTAAATCAGAGAAAGCCCAACTTT---GGATATGCAAGTCTCTGGAGAAAGAGAGA	10793
Qy	299	Thr-SerPheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlu	318

Db	10794	ACTAGAGGCTAGTGTGCACAAAAGCAAGGTGTATCAGGATGTCCAAAAGAAGTCCCGAGAAACC	10855
Qy	318	nGlu-----AlaGluLysValAspGluGlu-----	326
Db	10854	AGAAGTACTGAAGTCATAGCTGTAGAAGATATCTGAAGAAAGATTGAAGAACCAAGAA	10913
Qy	327	----GluLysGluLysThrGluProAlaSerGlu-----	336
Db	10914	ACCTGAAGAAAGGAAACCTGAGGTTAAATCATCAGAAAGACTACAGTCTTGGATAGCAAGT	10973
Qy	337	----GluGlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyrGlu--	354
Db	10974	CCTGAAGAAAGAAAGAACTAGAGGCTGTAGTCACAAAGCAAGGTCATCAGGATGTGCAAAA	11033
Qy	355	-----LysVal-----	356
Db	11034	GAAGTCCAGAAAACAGAAAGTATCTGAGGTCGTAGCTGAGAAGGTATCTGAAGGAAAGAT	11093
Qy	356	IgluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAl	376
Db	11094	TGAAGAACCAAGAAACCTGGAAGTAAAGAAAGAACTGAGGTTAAATCA---GAGAAAGCCAC	11150
Qy	376	aProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValValAlaGl	396
Db	11151	AACTTTGATTAAGCAAGTCCTGGAAGAAAA--	11181
Qy	396	uValHisValSerThrValGluLysThrGluGluGlnGlyGly---GlyGlyGluAl	415
Db	11182	-----GAAGTGTAGGCTGTAGTCACAAAGCAAGGTCATCAGGATGGAAAAATC	11228
Qy	415	aGluGlyGlyValVal-----ValGluGlyThrGl	425
Db	11229	TCGAGATCATATTATAAAACATTCAAAGAAAGGCTCACAGAAATTATCCAGGCCATTAGG	11288
Qy	425	yGluSerLeuProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluPr	445
Db	11289	ATCTAGTCTC---GATGAAATTTTGAGAGAAAGTCGGGAAATTTGTCAATAACCTTGAAGA	11345
Qy	445	o-----AlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGlyAs	461
Db	11346	TGATAAGGTAGTGCCAAACATTTATTCAAATTAAGAGACCATATAGTGTACACCTATGA	11405
Qy	461	phisThrGln-----Le	465
Db	11406	CGGAAAAGAGAGAGAGAAABATAAGAAAAGAACTCTTCGAATCCTTTATTGAATGCTT	11465
Qy	465	uThrAspLeuSerProGlu-----	471
Db	11466	GTGTGAGGCTCCCTCGAAGCAGCTGAGAAAGTAAAACTCAACTATCTTAAGGAAATTA	11525
Qy	471	-----	471
Db	11526	AACCAATGTGATTTTAAACAAAAGCTACTATTCAACTTATCGACGACTCTAACTGTTCCAC	11585
Qy	471	-----	471
Db	11586	CAACCGCTCTTATTAATATCCCAACCTTTTAAATTTGGAGAGAGTTGCCGTTAAATTC	11645
Qy	472	----GluLysThrLeuProLysHisProGluGlyLeValSer-----	484
Db	11646	ATCAGAAACATATGTTGTATAAGTCATCGGAAAAAATGATTAGTTGCAACAGAGCTTTAT	11705
Qy	484	-----	484
Db	11706	GGATATTTTCGTAATCTTGTATGACTTTTATAGATCAGCAGACGGAAGTTTTTAAACCCAAA	11765
Qy	485	-----GluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySe	500
Db	11766	AATTGAAAATATTAAACACAGCTTCTCAGTGACTATGATTATATTGAAAAAAGATCG	11825
Qy	500	rProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysGlnLy	520

Db 11826 ACCCTTTAACGGCTGTAATGAAAGATTAATGTTGATCC----- 11871
 QY 520 sGlySarArgIyGIyGIyAspGluGluProGlyGluIleHisThr-- 539
 Db 11872 -----CAACATATCTTAATCTAT 11888
 QY 540 -----GluSerProGluSerAlaAspGlu---GlnIySGlyGluSerSe 553
 Db 11889 TATTGAGAAAGTTAACTGATTAATCATGATCAAAAGAAAAAGAGCTGCCAA 11948
 QY 553 rAlaSerSerProGluGluProGluGluThrThrCysLeuGluIySGlyProLeuGluAl 573
 Db 11949 CCGGAGAGCGGATAATTCTCCGACGAGAAAAGCA-----GAAAGAAAGCAAAAGAGAGA 12002
 QY 573 aProGlnAspGIyGluAlaGluGluGlyThrThrSerAspGIyGluIySGlySarArgIyGlu 593
 Db 12003 GATTAAAGACCTCTGAAGCAACACAGAAATCGAAAGTTTCGGAGAAAGAAATCA----- 12057
 QY 593 yIleThrProTrpAlaSerPheIySGlyMetValThrProIySGlySarValArgArgPr 613
 Db 12058 -----ATCGAAGAAAGAACTTGAAGCAAGAAAGAAAGAAAGCAAGACAGTC 12104
 QY 613 oSerGluSerAspIySGlyGluGluGluIySGlyValIySGlySerAlaThrLeuSerSerTh 633
 Db 12105 TGCTATTGATGAAATCCCAAAAGGCTGAGTGTCCGAAATCTGTCCGAAAAAGATTAC 12164
 QY 633 rAspSerThrValSerGluMetGln---AspGIyValIySGlyThrValIyGIyGluGluIy 652
 Db 12165 TCATGAGAAAGCAACAAGAAAGCAAAAGAGAGAGGTTAAG-----GATTTCGAAGCCAA 12218
 QY 652 sProGluGluProIySGlySarArgValAspThrSerValSerTrpGluAlaLeuIleGlySa 672
 Db 12219 GCCTTAAGAAAGCAAGCTCTTGAGAAAGAAATCAATCGAAGAAAGAAAGAACTT----- 12270
 QY 672 lGlySerSerIySGlySarArgAlaArgGlyAlaSerSerAspAspAspGIyGIyGlyProAr 692
 Db 12271 -----GAAAGCAAGAAAGAAAGAAAGCAAGCGAGTGTCTATTGATGAA----- 12312
 QY 692 gThrLeuGIyGIyAspSerHisArgAlaGluGluAlaSerIyAspIySGlyGluAlaGlyTh 712
 Db 12313 -----AAATCCCAAAAGCTGAGGTGTCCGAAATGTGTCCGAAAAAGATTAC 12359
 QY 712 rAspAlaValProAlaSerThrGlnGluGluAspGlnAlaGlnIySGlySerSerProGl 732
 Db 12360 TCATGAGAAAGCAACAAGAAAGCAAAAGAGGTTAAGATTCTGAAGCCAAAGCTTAA 12419
 QY 732 uProAlaGlySerProSerGluGluGlyValSerThrTrpGluSerPheIySarArgLe 752
 Db 12420 GAAGCA-----AAGGT 12431
 QY 752 uValThrProArgIySGlySerIySGlyLeuGluGluIySGlyAlaGlu-----As 769
 Db 12432 CTTGGAAGAAATCAATCGAAGAGAGAAACTTGAAGCAAGAAAGAAAGCAAGCGAG 12491
 QY 769 pSerSerValGluGlnLeuSerThrGluIleGluProSerArgIyGluGluSerTrpValSe 789
 Db 12492 CTCTGCTATTGATGAATAATCCCAAAAGCTGAGTGTCCGAAATGTGTCTCC----- 12543
 QY 789 rIleIySGlyPheIleProGlyArgArgIySGlySarArgAlaAspGIySGlyGluGluGlnAl 809
 Db 12544 -----GAAAAATTACTGATGAGAGAGAGCAAGCAAGAAAG 12575
 QY 809 aThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAlaVa 829
 Db 12576 TCAAAAGAGAGAG-----CTTAAGATTCTGAAGCCAAAGCTTGAAGAGCAAAAGGT 12626
 QY 829 lValProLeuSerGlyTrpAsnAlaValAlaGluArgGluIySGlyMetGluAlaGlnIyAsnTh 849
 Db 12627 CTTG-----GAGAAAGAAATCAATCGAAGAGAGAAAGAAAGCTTGAAGAGCAAGAAAGAA 12677
 QY 849 rGluLeuProGlnLeuGluIyValAlaValIyValSerGluGluGluLeuSerIySGlyThrLeu 869
 Db 12678 ACAGACGAGGCTGTCTATT-----GATGAAATAATCCCAAAAGGCTG- 12718

QY 869 aHisThrValSerValAlaValIleAspGIyThrArgAlaValThrSerValGluGluAl 889
 Db 12719 -----A 12719
 QY 889 rGserProSerTrpIleSerAlaSerValThr-----GluProL 902
 Db 12720 GGTGCTCGAATATGTCTCCGAAAAAGATTACTGATGAGAGGCAAGCAAGAAAGCCAAAAAG 12779
 QY 902 euGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluIySGlyPheI 922
 Db 12780 AGGAGGTTAAGGAGTTCTGAAGCCAAAGCTTAAGAGGCAAAAGCTTGGAGAAAGAAATCA 12839
 QY 922 lValGluGluThrProValLeuThrGlnThrLeuProGluGlyIySGlyAspAlaHisAspA 942
 Db 12840 TCGAAGAGAG-----AAACTTGAAGAAAGCAAGAAAGAAAGCAAGAGCG 12881
 QY 942 sMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerG 962
 Db 12882 AC-----TCTGCTATTGAT-----GAAAAATCCC 12905
 QY 962 lu-AlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMet 981
 Db 12906 AAAAGCTTGAAGGTGTCCGAAATGTCT-----TCGAAAAAGATTACTGATGAG 12953
 QY 982 ValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProVal 1001
 Db 12954 AAGGCAAGAAAGCTTGAAGAGAGGTTAAGATTCTGAAGCCAAAGCTTGAAGAAAGCA 13013
 QY 1002 GlnGluValGluSerGlyValLeuAspThrGlu-----GluGluGluArgGlnThrGln 1019
 Db 13014 AAGGCTTGGAGAAAGAAATCAATCGAAGAGAAAGAACTTGAAGAGAAAGAAAGCAAG 13073
 QY 1020 AlaIleLeuGlnAlaValAlaAspIySGlyValIySGlyGluSerGlnValProAlaThrGln 1039
 Db 13074 ACG-----GACTGTGCTATTGATTAATAATCCCAAAAGCTGAGGTGTCCGAAATGTCT 13127
 QY 1040 ThrValGlnArgThrGlySerIySGlyAlaLeuGluIySGlyValGluGluValGluAspSer 1059
 Db 13128 CCGAAAGAAATTAATCTGATGAGAGGCAAGAAAGCTTGAAGAGAAAGAGAGGTTAAGATTCT 13187
 QY 1060 GluValLeuAlaSerGluIySGlyGluIySGlyAspValMetProIySGlyProValGlnGluAla 1079
 Db 13188 GAA-----GCCAAGCTTGAAGAGCAAAAGCTTGGAGAAAGAAATCAATCGAAGAGCG 13241
 QY 1080 GluAlaGluHisLeuAlaGluGlySerGluThrGlnAlaThrProGluSerLeuGlu 1099
 Db 13242 AAATCTGAAGCAAGAAAGAA-----ACACAGAGCGAGTGTCTGATGAGAAATCCCAA 13298
 QY 1100 ValProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleIySGlyLeuGln 1119
 Db 13299 AAGGCTGAGGTGTCCGAAATGTCTCCGAAAAAGATTACTGATGAG-----AAGGCACAA 13352
 QY 1120 GlnLeuMetGluGlnAlaVal-----AlaProGluSerSerGluThrLeu 1134
 Db 13353 GAAAGCCAAAGAAAGAGGTTAAGGTTCTTAAGCCAAAGCTTGAAGAAAGCAAGAGGCTTGG 13412
 QY 1135 ThrAspSerGluThrAsnGIySGlySerThrProLeuAlaAspSerThrAlaAspGIyThr 1154
 Db 13413 -----GAGAA-AGAAATCAATCGAAGAGAAAGAAAGAACTTGAAGCCAAAGAAAGAA 13462
 QY 1155 GlnGlnAspGluThrIleAspSerGlnAspSerIySGlyAla-----ThrAlaAlaValArgG 1173
 Db 13463 CAGACGAGAGTGTCTATTGATGAATAATCCCAAAAGGCTTGAAGGTGTCCGAAATGTCTC 13522
 QY 1173 nSerGlnValThrGluGluGluAlaAlaThrAlaGlnIySGlyGlu-----ProSe 1190
 Db 13523 CGAAAGATTACTGATGAGAGGCAAGAAAGCCCAAGAGAGGTTAAGGATTCTGA 13582
 QY 1190 rThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGlyArgAspVal 1210
 Db 13583 ACCCAAGCTTGAAGAGCAAGAGCTTGTGAGAAAGAAATCAATCGAA-----GAGAGAGA 13636

QY 1210 lLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAlaLysThrGI 1230
 Db 13637 ACTTGAAACAAAGAAAGAAAGACGAGCTCTGCTATT-----GATGA 13681
 QY 1230 uValGlyGlnGluGluValAspTrpLeuAspGlyGluLysValLysGluGluGlnGI 1250
 Db 13682 AAAATCCAAAGGCTGAGGTGCCGAACTGTCTCCGAAAGATTACTGATGAGAAGGC 13741
 QY 1250 uValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTyrAspSerGI 1270
 Db 13742 A-----GAAGAAAGCCGAAAGGAGGAGGTAAAGGACTCTGA 13777
 QY 1270 uValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSerLeuSerLe 1290
 Db 13778 A-----GCCAAGCTTAAGAAAGCAAGGCTCTTGAGAAAGAAATCAAT 13819
 QY 1290 uGluGluGlyGluMetGluThrAspValGluLys----- 1301
 Db 13820 CGAAGAGAGAACTTGAAGACAGAAAGAAACACAGACAGACTCTCTATTGATGAAAA 13879
 QY 1302 -----GluLysArgGluThrLysProGluGlnValSerGluGluGluGI 1317
 Db 13880 ATCCCAAAAGGCTGAGGTGCCGAACTGTCTCCGAAAGATTACTGATGAGAAGGCACA 13939
 QY 1317 nGluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMe 1337
 Db 13940 AGAAGCCAAAGAGAGGTTAAGGATTCTGAAGCCAGCCT----- 13982
 QY 1337 tProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAs 1357
 Db 13982 ----- 13982
 QY 1357 pGlnAspLysAlaGlyCysLeGlu-----ValGlnValGlnSerLeuAspThrTh 1374
 Db 13983 -----AAGAAAGCAAGATTTGGAGAGAAATCAATCGAAATAGAGAAATTTGGAT 14033
 QY 1374 rValThrGlnThrAlaGluAlaValGluLysValIleGluThrValValIleSerGluTh 1394
 Db 14034 -----GAAAGAAAGAAAGAAAGCAGCAGCAGACTAAGGTTGCCACTGATAC 14077
 QY 1394 rGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaTh 1414
 Db 14078 C-----AAATCCCAACGGTTGAGTTCCGAAATCGTTTGGAAAGATATCGGAA-- 14129
 QY 1414 rGlyGlyHisTrpThrLeuGlnHisAlaGluAspThr-----ValProLeuGlyProGI 1432
 Db 14130 -----GAAAGGCGAGAGAAAGCAAAAGGTTGAACTT---AAGGA 14167
 QY 1432 uSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeuHisPr 1452
 Db 14168 CTCTGAGGCCAAATCAAAAGCAAGCAAGGTCCTG---GAAAGAAAGTCAACCTTGAAGA 14224
 QY 1452 oAspLeuGlnGlyLysSerAlaSerGlnArgGlu-----ArgSerGI 1467
 Db 14225 GAACCTCGAT---GAGATGATATAAAACAGAGAGGATGCGCTACTAACAAATCCCA 14281
 QY 1467 uGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAla----- 1485
 Db 14282 AAAGCGAGGCTGCTGATCTGTGTCCGAAAAAATATCTGAAGAGAGGTTGCTGAAT 14341
 QY 1486 -----IleGluLysValLeuLysAlaGluProGluLeuLeuGI 1500
 Db 14342 AAAAATCTCTGAACCTTGAACCTCAAGGCTAAATCTAAACACAGATTACTCTGCCGA 14401
 QY 1500 uSerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaAr 1520
 Db 14402 TGAAGAGGTCACGGGCTAAGGTTCTCGAATCTGTTCTGTTAAAGAAATGAAGCTGAGAA 14461
 QY 1520 gThrGlu-----ThrAlaProGluThrHisAlaTyAspSerGlnThrGlnValProAl 1538
 Db 14462 AACAGATCAACTTTCAGCAAGAAACCAACTGTCTCGATGAGGATCTGGTTGTTCCAA 14521
 QY 1538 aCysArg----- 1540

Db 14522 ACAGAAAGCCATATTTCGGCTGAACAAACTGCAGATTCTATTAGCTCCTCAACAGTACAAGTC 14581
 QY 1541 -LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysMetLysAspAlaLysMetLysHi 1560
 Db 14582 CATGGACTCCGAGTACAAGGACCGA-----AAGGAATCCGGAAGTGCCTCAAGCGGAAG-- 14633
 QY 1560 sProValProGlnProArgGluAspLeuGlnVal 1571
 Db 14634 -----CCACCGTTGACATCCAACCTA 14654

RESULT 15

AAS80823

ID AAS80823 standard; cDNA; 7568 BP.

XX

AC AAS80823;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #16627.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

PR

23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI: 2001-639362/73.

DR

P-PSDB; ABG16636.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX

PS Claim 1; SEQ ID No 16627; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 7568 BP; 2332 A; 1873 C; 1825 G; 1538 T; 0 other;

Alignment Scores:

Pred. No.:	3,286-16	Length:	7568
Score:	492.00	Matched:	386
Percent Similarity:	33.47%	Conservative:	269
Best Local Similarity:	19.72%	Mismatches:	714
Query Match:	6.09%	Indels:	589
DB:	23	Gaps:	75

US-09-902-432-4 (1-1596) x AAS80823 (1-7568)

```

Qy 45 AlaAspProAlaThrLysLeu-----ProGlnLysAsnGlnGlnLeuSer 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1593 GCAAACCTCGCGAGAAATCATCCAGTCTGTTCTTCGGAAACGACACCAAGTCAAC 1652
Qy 60 SerValAsnGlyValAlaGlnGlnGlnLysAspValHis-----ValGln 73
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1653 ATCTCGAAGAGGTTGGAAAGCTCAAAACATCTAGACTTCTGAAGCAGCCAGTGCACC 1712
Qy 74 GlnGlnAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 93
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1713 CAAAAGGATCTCAGTCCAGGCTCCACTCTCTGTGTGAACAAACAAACTGAAGACG 1772
Qy 94 SerGlnAspValArgGlnLysAspArgValGlnGlnMetAlaAlaAsnSerThrAlaVal 113
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1773 AGGCTGATAGCCGAGAAAGT-----CTGAAGCCAGCCGCAAAACCACTCTTAC 1823
Qy 114 GlnAspIleThrLysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 133
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1824 AATCCGTGCGAGAGGATCAAAAGAAAGAACCCCTGAGGTC-----ACA 1868
Qy 134 SerGlnAsnAsnValGlnGlnMetValGlnProAlaGlnGlnGlnGlnGlnGln 153
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1869 AAGTCAATCAGCTGGAA-----AAGCCACCCCAAGTTGAAAGCAAGAA----- 1913
Qy 154 PheLysLysValPheLysPheValGlnPheLysPheThrValLysLysAspLysAsnGln 173
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1914 -----AAGTAAATGTTGAAAGAAAGCAAGCAAGTAA 1943
Qy 174 LysSerAspThrValGlnLeuLeuThrValLysLysAspGlnGlnGlnGlnGlnGln 193
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1944 AAAAAAGACCAACCACTTCACTGACTGAAGAGAG----- 1979
Qy 194 SerValGlyAlaGlyAspHisGlnGlnProSer---ValGlnThrAlaValGlnGlnSer 212
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1980 -----GTTCCACGCAAGAGAGCCATCTCCAGTAAAGCCAGGCTGTAG--- 2027
Qy 213 AlaSerLysGlnSerGlnLeuLysGlnSerThrGlnLysGlnGlnGlnGlnGlnGln 232
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2028 -----AAGCAAGCCACAGATGTCAAACCCAAAGCTGCCAAGAG 2066
Qy 233 GlnGlnSerSerThrGlnIleProLeuGlnAlaGlnSerAspGlnAlaAlaGlnGlnGln 252
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2067 AAGACGCGTAAAAAGGAAACAAAGGTAAGCCTGAAGACAAAGAAAGAGGAAAGAAAG 2126
Qy 253 AlaLysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 272
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2127 CCAAGGAAAGAGTGGCTAAAAAGGAGCAAAACATATCAAGAGAGGAAAGAAAGCA 2186
Qy 273 SerSerProValAsnSerGlnThrThrSerSerPheLysLysPheSerHisGlyTyr 292
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2187 AAA-----AAGGAAGAGGTGAAAAAGAAAGTCAAAAAAGAGTCTC----- 2225
Qy 293 AlaGlyTyrArgLysLysThrSerPheLysLysSerLysGlnAspLysLeuGlnThrAla 312
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2226 -----AAGAAAGAGAGGAAAAAGAAAGCAAGAAAGAGGTAAAGAAAGAAAGACCG 2276
Qy 313 GlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 332
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2277 CCAAGGAGAGTCAAGAGAGAGTAAAGAG-----GAGAGAGAGAGAGAG 2321
Qy 333 ProAlaSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 352
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2322 GTGAAAAAGAGAAAGAGAAAGCAAAAGAAAGATT-----AAGAGAGCTCCTTAAAGAC 2375

```

```

Qy 353 TyrGlnLysValGlnLeuProLeuGlnAspGlnValGlnAspLeuGlnAlaSerSerGln 372
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2376 GCAAAGAAATCATCTACCTCTCTG-----TCTTAAGCA 2408
Qy 373 GlnLysCysAlaProLeuAlaThrGlnValPheAspGlnLysMetGlnAlaHisGlnGln 392
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2409 AAAAAACAGCTGCTTTTAAAAACCAAAAGTACCAGAAAGAGAGTCTGTCAAGAAAGAT 2468
Qy 393 ValValAlaGlnValHisValSerThrValGlnLysThrGlnGlnGlnGlnGlnGln 412
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2469 TCTGTTGCTGCCGGAAGCCAAAGAGAGAGGAAATTAAGATCAATTAAGAGAGAGGC 2528
Qy 413 GlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 432
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2529 AAGCGCGAGGCGTGTGCTGCGACCTGTCCGACTGGAGGCAACAGCAGCTGTCTATG 2588
Qy 433 LeuAlaGlnProGlnGlnValProGlnGlnAlaGlnProAlaGlnGlnGlnGlnGlnGln 452
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2589 GCGGAGCTGGATAGCAGCC-----ATTGACCTGCGCAAGAACTGAGACTGAG 2639
Qy 453 ArgGlnMetCysValSerGlyGlyAspHisThrGln---LeuThrAspLeuSerProGln 471
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2640 AGGTCCCTTATGTATCTCTCTGAGAGATCTTAACCAAGACTTTGAAGAGTTAAGGCTGA 2699
Qy 472 GlnLysThrLeuProLysHisProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 491
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2700 GAGTCGATGTAAACAAAG-----GACATCAAGCTCTGAGCTGAGCTAATCAAGAC 2750
Qy 492 GlnGlnLysGlyLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 511
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2751 GAAGGAAACCTGAAG---GAATCTGAGCTCAAGCTTACGTCATCCAGAGAGAGA 2807
Qy 512 LysLysLeuSerGlyLysLysGlnLysGlnLysArgGly----- 524
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2808 GAAGTCACCAAGAGTCCAGCCAGTCCCTGATGAGGAATCATTACACACTGAAGGGAG 2867
Qy 525 GlyGlyLysAspGlnGlnGlnProGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 544
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2868 GCGGATGTGAACACACACTGAGAGCTGAGCGCCGTGCAAGACAGAGTAAAC--- 2924
Qy 545 AlaAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 564
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2925 ---GACATTGAAAAATTTGAAGATGAAGAGCGGTTTGAAGATCTTCAGAGACT--- 2978
Qy 565 CysLeuGlnLysGlyProLeuGlnAlaProGlnAspGlyGlnAlaGlnGlnGlnGln 584
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2979 -----GAGACTATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3026
Qy 585 SerAspGlyGlnLysLysArgGlnGlyIleThrProThrAlaSerPheLysLysMetVal 604
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3027 GAGGATGGGAGAGACAC-----GTA 3047
Qy 605 ThrProLysArgValArgArgProSerGlnSerAspLysGlnGlnGlnGlnGlnGln 623
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3048 TGTGTGAGCGCTCCAGAGACAGCCCACTGAGATGAGAAAGTGCACAGGCGGAGGCT 3107
Qy 624 -----LysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGln 639
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3108 GATGATACATGAGGAGAGAGAGAGAGAGTCTGTGCGCAGTGGGAGATGACCGAGCAAGA 3167
Qy 640 MetGlnAspGlnValLysThrValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 659
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3168 GACATGATGAGCGCATTTGAGAAAGAGAGCGCTGAACATCTGAAGAGAGCGCTGATGAG 3227
Qy 660 ValAspThrSer-----ValSerTyrGlnAlaLeu 669
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3228 GAGGACAAAGCTGAAGATGCCAGAGAGAGAGAAATATGAGCCGGAAGAAATGGAACCTGAA 3287
Qy 670 IleCysValGlySerSerLysArgAlaArgLysAlaSerSerSerAspAspGlnGly 689
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3288 GACTATGTATGTGCTGTGTGTCAGCAAGGCTGCAAGGCTGTGTGTGCGAGAGAGCATGAT 3347

```

Qy	690	Gly	-----ProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAla	704
Db	3348	GGATTCCTCACCACCAACCAAGCACTAGGAGCCCACTCTCTCTGGCCGAGAACTCTGCA	3407	
Qy	705	SerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGln	724	
Db	3408	TCITCAATTCATGAT-----GAGACTTACCTGGAGGCTCAGAGGCGAGGCCCAACC	3458	
Qy	725	AlaGlnGlySerSerProGluProAlaGlySerProSerGluGlyGluGlyValSer	744	
Db	3459	GCTTCTCATCAGGAGAATCGAGAA-----GACCAGCCTGAGGAATTCATCTGCCACCTCT	3512	
Qy	745	ThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGlu	764	
Db	3513	GGCTAC-----ACTCAGTCTACTATTGAG	3536	
Qy	765	GluLysAlaGluAspSerSerValGluGlnLeuSerThr	777	
Db	3537	ATATCCAGTAGGCCACCCCATCGGATGAGATGTTTCCCTCGAGAGCTGATGATGAT	3596	
Qy	778	-----GluLeuGluProSerArgGluGluSerTrpValSerLysLysLys	792	
Db	3597	GAGACCAACAATGAGAGAGCGGAGTCCCTCTCTCAGGAATTCGTAAATATACCAAAATAT	3656	
Qy	792	-----	792	
Db	3657	GAATCTTTCATTGTATTCTCAGGAATACTCTAAACCTGCTGATGTTACACCGCTCAACGGA	3716	
Qy	793	PheileProGlyArgArgLysArgAlaAspGlyLysGlnGluAlaThrValGlu	812	
Db	3717	TTTTCTGAAGGATCAAAAACAGATGCCACTGATGCAAGGATTACAATGCTTCAGCCTCT	3776	
Qy	813	AspSerGlyProValGluLeuAsnGluAspProAsnValProAlaValProLeu	832	
Db	3777	ACCATATCACCACCC-----	3791	
Qy	833	SerGluTrpAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGluLeuPro	852	
Db	3792	-----TCTTCCATGGAGGAACAATAATTCAGCAGATCTGCT-----	3827	
Qy	853	GlnLeuLeuGlyAlaValTrpValSerGluGluLeuSerLysThrLeuValHisThrVal	872	
Db	3828	-----TTACGTGATGCTTACTGCTCTGAAGTGAAGCCAGCACCC-----	3866	
Qy	873	SerValAlaValIleAspGlyThrArgAlaValThrSerValGlu-----GluArg	889	
Db	3867	ACTTTGGACATCAAGATAGCATCTCAGCTGTTTCAAGTGAAGAGTCAGCCCATCGAAG	3926	
Qy	890	SerProSerTrpLysSerAlaSerValThrGluProLeuGluHisThr-----AlaGlyGlu	908	
Db	3927	AGCCCGTCC---CTGAGTCCATCTCCACCATCACCTTTAGAAAAAGACCCCTGGGTGA	3983	
Qy	909	-----AlaMetProProValGlu-----GluValThrGluLysAspIlelle	922	
Db	3984	CGTAGTGTGAACCTTCTCTCGACGCCCATGAGATTAAGTCTCTGACAGGCGCAGAGTA	4043	
Qy	923	AlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAsp	942	
Db	4044	GCCCCGTGTCTCTCCTGAGGTGACCCAGAGTAGTTGGAACAATTTGTGTAGTCTCTGAG	4103	
Qy	943	MetValThrSerGluValAspPheThrSerGluAlaValThrAlaThr-----	958	
Db	4104	GACAAGACTCTGGAAGTGTGTGCCATCTCAGTCCGTGACTGGCAGTGTGTCACACA	4163	
Qy	959	-----GluThrSerGluAlaLeuArgThrGluGluVal-----	969	
Db	4164	CCTTACTATCAATCTCTACTGACGAGAAATCCAGTCTATCTCCCTACAGAAGTCATTGA	4223	
Qy	969	-----	969	
Db	4224	AAACCACGACGTTCCAGTGATGTTTTGAATTCAGTGATGCCAAGATGAGATGAAGG	4283	
Qy	969	-----	969	

Db	4284	GCCTTCAGTAAGCCCAATGGATGAGCCGTCCTGACTCAGAGTCTCTCTATTGAAAGAGTT	4344
Qy	969	-----	969
Db	4344	TTGTCTCCTTTACGACGCCGCCCTCATTTGGATCCGAGTCTGCTTATGAAAGTTTTCTA	4403
Qy	970	-----	974
Db	4404	AGTCCTGATGACAAAGCTTTCTGGCAGAGGTGCCGAAAGTCCCTTTTGAGAAAGAGAGTGA	4463
Qy	975	AlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSer-	991
Db	4464	AAACAAGGCTCTCCAGACCAAGTAAGTCCAGTTTCTGAAATGACTTCTACTAGTCTTTAC	4523
Qy	992	ProAspThrThrGlu-----GluAlaThrProValGlnGluValGluSerGly	1007
Db	4524	CAAGACAAACGAGGAGGAAAGACACAGACTTTGCACCAATAAAGAAGACACTTTGGCCAA	4583
Qy	1008	ValLeuAspThrGluGluGluArgGlnThrGlnAlaLeuLeuGlnAlaValAlaAsp	1027
Db	4584	GAAAGAAGAAACTGATGATGTTGAAGCCATGAGTTCTCAACGAGCACTGGGCTCTGGTGA	4643
Qy	1028	Lys-----	1028
Db	4644	AGGAAATTAGAGATGTTTCTCCACACAAATAGATGTCAGTCAGTTTGGATCTTTTAAA	4703
Qy	1029	-----	1037
Db	4704	GAAGACACTAAGATGTCCTATTTCTGAAGTACTGTCTCAGACCAAGTCAGTACTCTCTGT	4763
Qy	1038	ThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluValGluGlu	1057
Db	4764	GATGAGGCGTCACAGACAGACAGTACTCTCATATGGAGGCTGGGCTCAGTGTCCACA	4823
Qy	1058	AspSerGluValLeuAlaSerGluLysGlu-----LysAspValMetProLysGly	1074
Db	4824	GCCTCAGTGGCTACGAGCTCATTTCCAGAGCCCAACACAGATGATGTCTCCATCTCTG	4883
Qy	1075	ProValGlnGluAlaGlyAlaGluHis-----	1083
Db	4884	CATGCT---GAGGTGGCTCCCCACATCTCCACAGAGTAGATGACTCCCTTTTCAGTGTCT	4940
Qy	1084	LeuAlaGlnGlySerGluThrGlyGlnAlaThr-----	1094
Db	4941	GTTGTGCAACACTACACATCTCAGGAACAGAAATGTCCTCATTAAGAAGAATGC	5000
Qy	1095	-----ProGluSerLeuGluValProGluVal-----	1104
Db	5001	CCAAGACCGATGTCATTTCTCCACAGATTTCTCCCTCTAACTGCAAAAGTCCAGGACA	5060
Qy	1105	AlaaspValAspHisValAlaThrCysGlnValLleLysLeu-----GlnGlnLeu	1121
Db	5061	CCCGTTCACAGATCACAGATCTGAACAGTCCTCAATGTCTATTGAATTGGCCCAAGATCT	5120
Qy	1122	MetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThr-----	1139
Db	5121	CCTGAGCAATCCCTTCATATGACTTCAGTCGACAGTCTCCAGATCACCTACAGTGGGT	5180
Qy	1140	-----AsnGlySerThrProLeu-----AlaAspSerAsp	1149
Db	5181	GCAGGCGTGCTTCACATCACTGAAATGGCCCACTGAAGTGGACTACAGTCTCTCTGAC	5240
Qy	1150	ThrAlaaspGlyThr-----GlnGlnAspGluThrLleAspSerGln	1163
Db	5241	ATGCAGGACTCCAGCTTTATCATTAAGATACCACTATGGAGGAGCGCTCCTACACCCAA	5300
Qy	1164	AspSertyAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThr	1183
Db	5301	GATAATCATCTTCTGAGCTCATCTCAGTATCTCAGGTA---GAGGCGTCCCGTCCACC	5357
Qy	1184	AlaGlnLysGluGluProSerThrLeuProAsnValProAlaGlnGluGluHisGly	1203

```

Db 5358 TCTTGTGCTATACCCCTTCTGAGTCGCTCT-----CCTCTCCA----- 5399
QY 1204 GTTGTGCTATACCCCTTCTGAGTCGCTCT-----CCTCTCCA----- 1223
Db 5400 GAAAGTACTCTATCCGATGTTGCTCTCCAGAGATATCTCTATATGCTCATCTACACC 5459
QY 1224 ProValLeuAlaLysThrGluValGlyGlnGluValAspThrLeuAspGlyGlu 1243
Db 5460 -----TCTGAAAAAGTGGAAAGTCTGGAAAGAGAG 5489
QY 1244 LysValLysGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAla 1263
Db 5490 AGCTCTCTCCAAATCTATATCTCTCCATCCACCCACGAGATCTCTCTTATAT 5549
QY 1264 AspValThrTyrAspSerGluValMetGlyVal-----AlaGlyCysGlnGlu 1279
Db 5550 TCACCTACTTTTCAATCTACCTCTGCAATCAAGAAAAACCAACTTGCCACAGT 5609
QY 1280 LysGluSerThrGluValGlnSerLeuSerLeuGluGluMetGluThrAspVal 1299
Db 5610 TCCTCTTCCCAATAGATGATCCGAGAGCCCTTATGCTTCCGTCCTCAGTG 5669
QY 1299 ----- 1299
Db 5670 TTATTGATACATGCAACACCATCTAGCTTGAATAGAGATTGTCCACACCTGGCTG 5729
QY 1300 GlnLysGluLysArgGluThrLysProGluGlnValSer-----GluGlnGluGlu 1316
Db 5730 GAAAGAGACAGTGGAGGAGAACACCTGGTGAATTAAGCTTATGCTATCAAAAGCTGAG 5789
QY 1317 GlnGluThrAlaAlaProGluHisGluGly-----ThrTyrGlyLysProVal 1332
Db 5790 GAAACCAACGAGTCCCAATGAAAGATGATGATGATGATGATGATGATGATGATGAT 5849
QY 1333 LeuThrLeuAspMet-----ProSer 1339
Db 5850 CGAGCTTCAGATGTGGTGGCTTATGATGAGATGAGAGAACCAAAATCTCCA 5909
QY 1340 SerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAsp 1359
Db 5910 AGTACAGAGTGGCTACTCTATGATGACCAATGGGAAAACTACCAAGACCCCTGAAGTGT 5969
QY 1360 LysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrValThrGlnThrAla 1379
Db 5970 -----GACTATTCCTAT 5981
QY 1380 GlnAlaValGluLysValIleGluThrVal-----ValIle 1391
Db 5982 GAAATTTATGAGAGACACACGAGCCCTGAAAGAGGCGGTACTCATATGACATAGT 6041
QY 1392 SerGluThrGlyLysSerProGluCysValGlyAlaHisLeuLeuProAlaGluLysSer 1411
Db 6042 GAAAGAACCAACGAGCCCGGAAAGTGGTGTACAGCTATGAAAAAGACTGAGAGTGT 6101
QY 1412 -----SerAlaThrGlyGlyHisTyr 1418
Db 6102 AGAAGCTTCTGATGATACATGACAAATGGCTATGATGATGATGATGATGATGATGAT 6158
QY 1419 ThrLeu-----GlnHisAlaGluAspThrValProLeuGlyProGlu 1432
Db 6159 ACACCTTGGGAGCCCAAGCTACTTATGAAACCACTGAGAAAAATTACAGTTTCCCTGAG 6218
QY 1433 SerGlnAlaGluSerIleProIleIle----- 1441
Db 6219 TCTGAAGGTATTTCTATGATGATCAACAAAGACAAAGAACCCCTATCTTCACACA 6278
QY 1442 -----ValThrProAlaProGluSerThrLeuHisProAsp 1453
Db 6279 TACTGTACAGAGCTGACAGAAATACATAGAACCTTCAGGATCCATCATATCTCTAC 6338
QY 1454 LeuGlnGlyGluLysSerAlaSerGlnArgGluArgSerGluGluAspLysProAsp 1473
Db 6339 GAGACTTCAGACTATGCTACTGCAAGAAAGAGTCCCTCAGAGAGCCGTCAGAGT 6398

```

```

QY 1474 AlaGly-----ProAspAlaAspGlyLysGlu 1482
Db 6399 GTGATTTATGCTGCTGCTGCTCTTGTGAATACAGACACCCCAAGACAGACTTTACACC 6458
QY 1483 Ser-ThrAlaIleGluLysValLeu----- 1490
Db 6459 TCTTTCATTAATCCATTCCTTGTGAGTGTTCGCAATGAAAGAACCCATGAAAGATCT 6518
QY 1491 -----LysAlaGluProGluIleLeuGluLeuGluSerLys 1504
Db 6519 GAAAGCCCTTCATCAATCAAGGGAGCCCAAGCTTCAGAGAGAAACAAACAGAGGC 6578
QY 1504 nLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaGlyThr-- 1523
Db 6579 CGACAGTGTGAAGAACCCCTCC--ACCTCAGTCAAGCAGAGTCCATCCACAGCCGA 6637
QY 1524 -----AlaProGluThrHisAlaTyrAspSerGlnThrGlnValProAlaCysAr 1540
Db 6638 CTCTGATGTTCCCGGAGACTGAAAGAGTCCCTTCATCAG-----GCCGATGCCAA 6691
QY 1540 GlnAspSerArgGluProAsnArgCysTyrThrLysMetLysAspAlaLysMetLysHis 1560
Db 6692 TATGACTGTGAAGACGAGTGGAAACATCCCAACAGACAAACTGTCAGTACAAACA 6751
QY 1560 S-----ProVal-----ProGlnProArgGlu----- 1567
Db 6752 CATGACCCACTCCAGCTCCCGTGCAGAACCGAGCCCTTCGACCAACCCCTGATGT 6811
QY 1568 -----AspLeuGlnValLeuThrValLeuGlu-AlaTyrAlaGln----- 1580
Db 6812 GTCCATGTGTGACCCAGAGGCTTGGCATTGAGCAGAACTGGGCAAAAGCTTAAAGA 6871
QY 1581 -----ProArgLysCysLeuProArgLeuGlnLeuLysAlaPro 1593
Db 6872 AGATCTGAAGAGAGAACCAAAACCAAAAGCCAGGTACAAAGACCA 6918

```

Search completed: December 13, 2002, 00:45:32
Job time : 1027 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 00:24:46 ; Search time 124 Seconds
(without alignments)
3947.227 Million cell updates/sec

Title: US-09-902-432-4

Perfect score: 8073

Sequence: 1 MGAGSTGSRSPQAGSDT.....AWAQRKCLPRLQLKAPVSK 1596

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US0902432/runat 12122002 141452 25831/app query.fasta 1.1735
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPTXT=0 -UNITS=BITS -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0902432 @C@N 1 1 66 @runat 12122002 141452 25831 -NCPU=6 -ICPU=3
-NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -PGAPOP=6 -PGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6005	74.4	5134	2	US-08-635-121-1
2	4281	53.0	6605	1	US-08-769-309A-4
3	4281	53.0	6605	3	US-08-994-570-4
4	462	5.7	5361	4	US-08-973-462-2
5	462	5.7	6152	4	US-08-973-462-1
6	421.5	5.2	6744	1	US-08-119-125A-2
7	415	5.1	6306	1	US-08-466-390-3
8	415	5.1	6306	1	US-08-470-950-3
9	415	5.1	6306	1	US-08-467-781-3
10	415	5.1	6306	2	US-08-483-924-3
11	412	5.1	6306	1	US-08-195-487-3
12	412	5.1	6306	5	PCT-US93-06160-3

13	410.5	5.1	6755	3	US-08-931-999-4	Sequence 4, Appli
14	388	4.8	11236	1	US-07-853-913-1	Sequence 1, Appli
15	366	4.5	9551	1	US-08-056-200-93	Sequence 93, Appli
16	366	4.5	9551	2	US-08-800-644-93	Sequence 93, Appli
17	365.5	4.5	6414	4	US-09-134-001C-1626	Sequence 1626, Ap
18	365.5	4.5	11091	4	US-09-134-001C-2243	Sequence 2243, Ap
19	357.5	4.4	7766	4	US-09-125-619-3	Sequence 3, Appli
20	354	4.4	8224	6	5180808-1	Patent No. 5180808
21	351	4.3	5661	4	US-08-938-105-2	Sequence 2, Appli
22	346.5	4.3	32207	2	US-08-770-379-20	Sequence 20, Appli
23	346.5	4.3	32207	4	US-08-757-669A-20	Sequence 20, Appli
24	346.5	4.3	32207	4	US-09-230-371A-20	Sequence 20, Appli
25	343.5	4.3	3337	1	US-08-072-610-1	Sequence 1, Appli
26	343.5	4.3	3337	2	US-08-719-822B-1	Sequence 1, Appli
27	343.5	4.3	3337	4	US-09-092-458-1	Sequence 1, Appli
28	339.5	4.2	3561	4	US-09-134-001C-1685	Sequence 1685, Ap
29	334.5	4.1	3489	2	US-08-728-323A-1	Sequence 1, Appli
30	334.5	4.1	3489	4	US-09-298-568-1	Sequence 1, Appli
31	333.5	4.1	4852	1	US-07-853-913-3	Sequence 3, Appli
32	321	4.0	5656	1	US-08-425-061-1	Sequence 1, Appli
33	321	4.0	5656	1	US-08-825-886-1	Sequence 3, Appli
34	321	4.0	5689	1	US-08-425-061-3	Sequence 3, Appli
35	321	4.0	5689	1	US-08-825-886-3	Sequence 3, Appli
36	321	4.0	5709	1	US-08-425-061-2	Sequence 2, Appli
37	321	4.0	5709	1	US-08-825-886-2	Sequence 2, Appli
38	321	4.0	5711	1	US-08-425-061-4	Sequence 4, Appli
39	321	4.0	5711	1	US-08-825-886-4	Sequence 4, Appli
40	321	4.0	5711	2	US-08-658-322-1	Sequence 1, Appli
41	321	4.0	5712	1	US-08-425-061-12	Sequence 12, Appli
42	321	4.0	5712	1	US-08-825-886-12	Sequence 12, Appli
43	321	4.0	5712	2	US-08-603-753D-1	Sequence 1, Appli
44	321	4.0	5712	3	US-09-009-753-1	Sequence 1, Appli
45	321	4.0	5712	4	US-08-986-106-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-635-121-1
; Sequence 1, Application US/08635121
; Patent No. 5910442
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,121
; FILING DATE: 19-APRIL-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30558 - 165/33603
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 5134 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 US-08-635-121-1

Alignment Scores:
 Pred. No.: 0 Length: 5134
 Score: 6005.00 Matches: 1232
 Percent Similarity: 97.33% Conservative: 7
 Best Local Similarity: 96.78% Mismatches: 27
 Query Match: 74.38% Indels: 14
 DB: 2 Gaps: 4

US-09-902-432-4 (1-1596) x. US-08-635-121-1 (1-5134)

QY 329 GtUyThrGluProAlaSerGluGluGluProAlaGluAspThrAspGlnAlaArg 348
 |||||
 Db 2 GAAAAGACAGAGCCAGCCCTCGAGAGACAGAGCCGGCAGAAACACACAGCCAGGCCAGG 61
 QY 349 LeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
 62 TTGTACGACGACTACAGAGAGGTGAGCTGCTTTGGAAGACCAAGTGTGTGACCTGGAG 121
 QY 369 AlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
 122 GGATCCTCGAGAGAGAGTGTCTCTTTGGCAACGAAAGTGTATATAGAAAGATGGAA 181
 QY 389 AlaHisGluValValAlaGluValHisValSerThrValGluLysThrGluGluGlu 408
 182 GCCCACCAGAAAGTTTTCAGAGGTCCACGTACGACCCGTGAGAAAGACAGAGAGAG 241
 Db 409 GtnglyGlyGlyGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 428
 242 CAGGGAGAGAGAGAGAGCTGAAAGGGGCGTGTGTGTAAGAAACAGAGAAATCTTG 301
 QY 429 ProProGluLysLeuAlaGluProGluGluValProGluGluAlaGluProAlaGluGlu 448
 302 CCCCTCGAAGAACTGGCTAGCCCCAGAGAGTCCCAGAGAACTGAGCTGTGAGAG 361
 QY 449 LeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeu 468
 362 CTGATGAAAGACAGAGAGATGTGTCTCTGAGAGAGACCACTCACTGACAGACCTA 421
 Db 469 SerProGluGluLysThrLeuProLysHisProGluGlyTLeuValSerGluValGluMet 488
 422 AGTCCTGAAAGAGAGCGTGCACCAACCCAGAGGCACTGTCTGAGGTGGAGATG 481
 QY 489 LeuSerSerGluGluArgLysValGluLysValGluLysSerProLeuLysLysLeuPheSer 508
 482 CTGTCTCTCAGAGAAAGATCAAGGTACAGGGAAAGTCCCTTGAAGAACTCTTCACTAC 541
 Db 509 SerGlyLeuLysLysLeuSerGlyLysLysGluLysGlyLysArgGlyGlyGlyAsp 528
 542 TCAAGGCTTAAAGAGCTGTCTGGAGAGAGACAGAGAGGAGAAAGAGAGGTGGGAGAG 601
 QY 529 GtUgluProGlyGluLysGluLysHisLeuHisThrGluSerProGluSerAlaAspGluGlu 548
 602 GAAAGACCTCGAGAAATACCAACATTCACACCGAATCCCAAGAGAGTGTATGAGCAG 661
 Db 549 LysGlyGluSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLys 568
 662 AAGGGAGAGAGCTCTGCTGCTCCCGAGAGAGCTGAGAGAGACCACTGTCTGGAGAAA 721
 QY 569 GtYProLeuGluAlaProGluAspGlyGluLysGluGluGlyThrThrSerAspGlyGlu 588
 722 GGGCGCTGTGAGAGACCCAG-GATGGGAAAGCTGAGAGAAAGAACTTCTTCTG--GGAGAG 778

QY 589 LysLysArgGluGlyTLeuThrProTPrAlaSerPheLysLysMetValThrProLysLys 608
 |||||
 Db 779 AAGAGAGAGAGG--ATCATTCCCTGGGCATCTTCAAAAAAGTGTATACCCCAAGAAA 836
 QY 609 ArgValArgArgProSerGluSerAspLysGluGluGluLysValLysSerAla 628
 837 CGGTC-CGAAAGCTTCTTGAGAGTACAGAGAGAGAGCTGAGAGAGGTCAAGAGCCG 895
 Db 629 ThrLeuSerSerThrAspSerThrValSerGluMetGluAspGluValLysThrValGly 648
 896 ACCTGTCTCCTCACTGATAGCAGAGTGCAGAAATGCAAGATGAAGTCAAAACTTGTGT 955
 QY 649 GtUgluGluLysProGluGluProLysArgValAspThrSerValSerTPrGluAla 668
 956 GAGGAACAAAAGCCAGAGAACCAAGGCTGAGTACTTCACTGTCTTGGAGAACCA 1015
 QY 669 LeuLysCysValGlySerSerLysLysArgLysArgLysAlaArgLysAlaSerSerAspAspGlu 688
 1016 CTGATTTGTGTGATCATCCAG 1075
 Db 689 GlyGlyProArgThrLeuGly-GlyAspSerHisArgLysGluLysAlaSerLysAspLys 708
 1076 ---GGGCCAAGGACCTGGAGGGGGACATGCACAGACAGAGAGAGGCGCAGAAAGACA 1132
 QY 708 GtUgluLysGlyThrAspAlaValProAlaSerThrGluGluLysAspGlnAlaGlnGlySe 728
 1133 AGAAGCCCA-ACAGAGCTGTCTTCCAGACACCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
 QY 728 rSerSerProGluProAlaGlyLysSerProSerGluGlyGluGlyValSerThrTPrGluSe 748
 1192 TTCTCACCACCGAG 1251
 Db 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGluVal-- 767
 1252 ATTTAAAGATTAGTCACTCCAGAAAATAATCCAAAGTCAAACTGGAAGAAAGAA--AG 1310
 QY 768 ----GluAspSerSerVal---GluGlnLeuSerThrGluLysGluProSerArgGluGlu 785
 1311 CCGAAGAGACTCTAATTGTAG 1370
 Db 785 uSerTPrValSerLysLysPheLysProGlyArgArgLysLysArgLysAspGlyLys 805
 1371 ATCTTGGGTTTCAATTAAAGAAATTCATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1430
 QY 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluLysLeuAspAspProAla 825
 1431 GGCAGAAACAAGCCACTGTGGAGAGCTCAGGGCCAGTGGAGATGAATGAGAGAGAGAGAG 1490
 Db 825 snValProAlaValValProLeuSerGluLysPheAlaValAlaGluArgGluLysMetGlu 845
 1491 ATGTCCAGAGCGCTCGTGCCTGTCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1550
 QY 845 lAgGlnGlyAspThrGluLeuProGluLysLeuGluValValTyrValSerGluGluLys 865
 1551 CCCAGGGAGATCGGAGAGCTGCGAGCTGCTGGGCTGTGTA---GTGTCCGAGAGGCTCA 1607
 Db 865 eLysThrLeuValHisThrValSerValAlaValLysAspGlyThrArgLysAlaValThr 885
 1608 GTAAAGCTCTGTCCACACTGTGAGTGTGAGATGATGAGAGAGAGAGAGAGAGAGAGAG 1667
 QY 885 eValGluGluLysSerProSerTPrLysSerAlaSerValThrGluProLeuGluLysHis 905
 1668 GTGTGGAAGAGGCTCTCTCTGTGATATCGGCTTCCGTAAACAACACTTGTGAACACA 1727
 Db 905 hAlaGlyGluAlaMetProProValGluGluValThrGluLysAspLysLeuLeuLeuGlu 925
 1728 CAGCGGAGAGAGCCATGACCTGTGAGAGAGTCACTGAAAGACATCATTTGAGAGAG 1787
 QY 925 lUThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945
 1788 AAACCTCTGTGTCAACCAAGAGCTTACCAAGAGAGGTAAAGATGCCATATACAGCATGTCTCA 1847

QY 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
 Db 1848 CCAGTGAAGTGGATTTCACCTCAGAGCTGTGACAGCCACAGAGACCTCAGAGCTCTCC 1907
 QY 965 rgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlav 985
 Db 1908 GTACTGAAGAAGTTACCGAAGCATCGGGGCCGAAGAGAGACCACAGACATGGTGTCCGCGAG 1967
 QY 985 alSerGlnLeuThrAspSerProaspThrThrGluGluAlaThrProValGlnGluValG 1005
 Db 1968 TTTTCCAGAGTGAAGTCTCCCGACACCCACAGAGGAAGCCACCCAGTTTCAGAGGTAG 2027
 QY 1005 luSerGlyValLeuAspThrGluGluGluGluGluGluGluGluGluGluGluGluGlu 1025
 Db 2028 AGGT 2087
 QY 1025 alAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
 Db 2088 TTGCAGACAAGGTGAAGAGGAGTCCAGGTGCTGCAACCCAGACGTGTGCAGAGAAGCG 2147
 QY 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
 Db 2148 GGTCAAAAGCACTGGAGAAGGTTGAGAGGTTAGAGGAGTCCGAAAGTGTGCTTCGG 2207
 QY 1065 lylLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085
 Db 2208 AGAAAGAGAAGAGGAGTATATCCGAAAGAGCCCGTGCAGGAAGCTGGAGCATCTTG 2267
 QY 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
 Db 2268 CACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCTTGAAGTTCCTGAAAGTACAG 2327
 QY 1105 laAspValAspHisValAlaThrCysGlnValLysLeuGlnGlnLeuMetGluGlnA 1125
 Db 2328 CAGATGTAGACCATGTCCGACAGTGCAGGTATCAAGCTCCAGCAGTGTGAGAGCAAGG 2387
 QY 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
 Db 2388 CCGTGGCCCTTGAGTCTATCCGAACCTTGACAGACAGTGTGAGACAAATGGAAAGCACTCCCT 2447
 QY 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAsps 1165
 Db 2448 TAGCAGATTACAGACTGCGAGATGGGACACAGCAGATGAACCATTCACAGCCAGGACA 2507
 QY 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaG 1185
 Db 2508 GTAAAGCCACTGCAGCTGTCCAGGCAGTCCAGAGTTCACAGAAAGAGCGGCTACTGCTC 2567
 QY 1185 lnlLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluG 1205
 Db 2568 AGAAAGAGGAGCCCTTGACACTACCTAATATGTTCAGGCCCCAGGAGAAACATGGGGAAG 2627
 QY 1205 luProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProv 1225
 Db 2628 AACCGGAAGAGATGTTCTTGAACCTACACAGCAGAGCTTGTCTGTCAGCCGTCGCCCG 2687
 QY 1225 alLeuAlaThrThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysV 1245
 Db 2688 TCTGGCAAAAGACTGAGGTGGTCAAGAGGTTGAGTGTGCTGTTGATGGAGAAAG 2747
 QY 1245 alLysGluGluGlnGluValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspV 1265
 Db 2748 TCAAAGAAAGAACAGGAGGTGTTGTATCACTCTGGACCCAAACAGTCAGAAAGGCTGCTGATG 2807
 QY 1265 alThrThrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
 Db 2808 TGACATATGACATGAAGTGAAGTGGAGTGGCGGTGTACGAAAGAGAGAGTACTGAAG 2867
 QY 1285 alGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
 Db 2868 TGCAGAGTCTTAGCTGGAGGAGGAGAGATGGAACTGACGTGTGAAAGAGAGAAAGG 2927
 QY 1305 luThrLysProGluGlnValSerGluGluGluGlnGluThrAlaAlaProGluHisG 1325

Db 2928 AGCAAAAGCCAGAGCAAGTGAAGAGGTGAGCAGGAAACAGCCGCTCTCTGAGCATG 2987
 QY 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluAtgGlyLysA 1345
 Db 2988 AAAGGAACATACGGGAAGCCAGTCTTGACACTTGACACTGCCAGCTCAGAGAGGGGAAGG 3047
 QY 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
 Db 3048 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCCAGACCAAGACAAAGCAGGTTGCATAG 3107
 QY 1365 luValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysV 1385
 Db 3108 AGGTTCAAGTTCAAGCTCGGACACACAGTCACTCAACAGCAGAGCTGTGGAAAGG 3167
 QY 1385 alIleGluThrValValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisL 1405
 Db 3168 TCATAGAACGGTGTGATTTCCAGACAGAGTGAAGTCCAGAGTGTGTAGGTGCACACT 3227
 QY 1405 euLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluA 1425
 Db 3228 TATTACAGCTCAGAAAGTCTCTGCAACGGGTGGCCACTGGACTTTCAGCATGCAGAGG 3287
 QY 1425 spThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProA 1445
 Db 3288 ACACGGTACCCCTGGGGCCCTGAGTCTCAGGCAGAAATCCATCCCAATCATAGTAACCTCTG 3347
 QY 1445 laProGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluA 1465
 Db 3348 CTCCTGAAAGCACCTTACATCTCGACCTACAGGAGAGAAATAGCGCATCCAGAGAGAGC 3407
 QY 1465 rGserGluGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThrA 1485
 Db 3408 GATCAGAGGAAGAGGACAAGCCAGATGCTGCTGCTGATGCTGACGCAAGGAGAGTACAG 3467
 QY 1485 laIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnL 1505
 Db 3468 CAATCACAAGAGTCTCAAGGCTGAACCTTGAGATCTCTGAACTTGAGAGTAAGAGCAACA 3527
 QY 1505 yeIleValLeuAenValIleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaP 1525
 Db 3528 AGATTGTCTGAACGTCATTCAGACAGCCGTTGACCACTTCGCACCTCAGCAAAACAGCCC 3587
 QY 1525 roGluThrHisAlaTrpAspSerGlnThrGlnValProAlaCysArgLeuAspSerArg 1545
 Db 3588 CCGAAACTCATGCTTATGATTCACAGCCAGGTTCTGCAATGCTGCTGGACAGCAGGG 3647
 QY 1545 luProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHisProValProGlnP 1565
 Db 3648 AGCCCAACAGATGCTGACCAAAATGAAAGTTGCCAAGATGAAACACCCAGTCCCGCAGC 3707
 QY 1565 roArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCysL 1585
 Db 3708 CCAGAGAGACTTGCAGTCTCTGACCGTCTTGAGGAGCATGGC--TCAGCTCGGAAATGCT 3765
 QY 1585 euProArgLeuGlnLysAlaProValSerLys 1596
 Db 3766 TGCCGCGCTTGCAAGTTGAAAGCGCCGCTGTCAAAG 3800

RESULT 2
 US-08-769-309A-4
 ; Sequence 4, Application US/08769309A
 ; Patent No. 5741890
 ; GENERAL INFORMATION:
 ; APPLICANT: Scott, John D.,
 ; APPLICANT: Nauert, Brian J.,
 ; APPLICANT: Klauck, Theresa M.
 ; TITLE OF INVENTION: Protein Binding Domains of Gravin
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower/233 South Wacker Drive
 ; CITY: Chicago

```

1      STATE: Illinois
2      COUNTRY: United States of America
3      ZIP: 60606-6402
4
5      COMPUTER READABLE FORM:
6
7      MEDIUM TYPE: Floppy disk
8      COMPUTER: IBM PC compatible
9      OPERATING SYSTEM: PC-DOS/MS-DOS
10     SOFTWARE: Patent Release #1.0, Version #1.30
11
12     CURRENT APPLICATION DATA:
13     APPLICATION NUMBER: US/08/769,309A
14
15     FILING DATE:
16
17     CLASSIFICATION: 435
18
19     ATTORNEY/AGENT INFORMATION:
20     NAME: No. 5741890and, Greta E.
21     REGISTRATION NUMBER: 35,302
22     REFERENCE/DOCKET NUMBER: 27666/33451
23
24     TELECOMMUNICATION INFORMATION:
25     TELEPHONE: 312-474-6300
26     TELEFAX: 312-474-0448
27
28     TELEX: 25-3856
29
30     INFORMATION FOR SEQ. ID NO.: 4:
31
32     SEQUENCE CHARACTERISTICS:
33     LENGTH: 6605 base pairs
34     TYPE: nucleic acid
35     STRANDEDNESS: single
36     TOPOLOGY: linear
37
38     MOLECULE TYPE: cDNA
39
40     FEATURE:
41
42     NAME/KEY: CDS
43     LOCATION: 192..5531
44
45     US-08-769-309A-4

```

Qy 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
Db 1707 GAGAAATGAGGTGCGAGGAAAGTCCACTAAAGAAAGCTTTTACCAAGCTGGCTTAAAA 1766
Qy 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyAspGluGluProGly 532
Db 1767 AAGCTTCTGGAAGAACAGAAAGGNAAGA---GGAGGAGGAGCAGGAGATCAGGG 1823
Qy 533 GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1824 GAGCACACTCAGGTTCACGCGGATCTCCGACAGCCAGGAGGACAAAGGGCGAGAGC 1883
Qy 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1884 TCTGCTCATCCCTGAGGAGCCCGAGAGATCACGTGCTTGGAAAGGGCTTAGCCGAG 1943
Qy 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1944 GTGCAGCAGGATGGGGAAGCTGAGAGAGGAGCTACTTCCGATGGAGAAAGAGAGAA 2003
Qy 593 GlyIleThrProTyrAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
Db 2004 GGTGTCACTCCCTGGGCATCATTCAAAGAGATGTGACGCCCAAGAACGCGTGTAGACGG 2063
Qy 613 ProSerGluSerAspLysGluGluGluLysLysValLysSerAlaThrLeuSerSer 632
Db 2064 CCTTCGAAAGTGAATAAGAGATGAGCTGCAAGGTCAAGAGCGCTACCTGTCTTCC 2123
Qy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluLys 652
Db 2124 ACCGAGACACAGCTCTGAAATGCAAGAAATGAAGAGGAGCGTGGAGAGCCAAAG 2183
Qy 653 ProGluGluProLysArgValAspThrSerValSerTyrGluAlaLeuIleCysVal 672
Db 2184 CCGAAGAACCAAGCCAGGTGATACCTCAGTATCTTGGGAAGCTTTAATTTGTGTG 2243
Qy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyProArg 692
Db 2244 GGATCATCCAAAGAAAGACGAGGAGAGGTCTCTTCTGATGAGGAGGGGACCAAAA 2303
Qy 693 ThrLeuGlyGlyAspSerHisArgAlaGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2304 GCAATGGAGGAGACCAACCAAGAAAGCTGTAGGCGCCGGAAGAAAGACAGCGGGACA 2363
Qy 713 AspAlaValProAlaSerThrGlnGlnAspGlnAlaGlnGlySerSerSerProGlu 732
Db 2364 GACGGATCCTCTGCTGTTCCCAAGACATGATCCAGGCGAGGAGAGTCTCTCCCGGAG 2423
Qy 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTyrGluSerPheLysArgLeu 752
Db 2424 CAAGCTGGAAGCCCTACCGAAGGGGAGGGCGTTTCCACCTGGGAGTCATTTAAAGGTTA 2483
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGluLysAlaGluAsp----- 769
Db 2484 GTACGCCAAGAAATAATCAAGTCCAGCTGGAAGAGAAAGCGAAGACTCCATAGCT 2543
Qy 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerTyrVal 788
Db 2544 GGGTCTGGTAGAACATTCACCTCCAGACACTGAACCCCGTAAAGAAAGATCTCTGGGTC 2603
Qy 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGln 808
Db 2604 TCAATCAAGAAAGTTTATTTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2663
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspProAsnValProAla 828
Db 2664 GCCCTGTGGAAGACCGAGGGGCCAACAGGGGCCAACCAAGATGACTCTGATGTCCCGGCC 2723
Qy 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2724 GTGGTCCCTCTGTCTGAGTATGCTGTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2783
Qy 846 GlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuSer 865

Db 2784 CAAAAAGCCGACAGAGCAGCGCCGAGAGAGAGCCACCTAGGTGTCCAAAGAGGCTCAGC 2843
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2844 GAGAGTCAGGTTCATATGATGGCAGCAGCTGTCTGCTGACGGAGCAGGCGAGCTACCAT 2903
Qy 886 ValGluGluArgSerProSerTyrIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2904 ATTGAAGAAGAGGTCTCTCTTCTGGATATCTGTTTCAGTCACAGAACTCTTTGAACA 2963
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
Db 2964 GAAGCTGAAGCCCACTGTTAACTGAGGAGGTATTGGAAGAGAGAGTAATTGCAAGAA 3023
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 3024 GAACCCCCACCGGTTACTTGAACCTCTGCCAGAGAACAGAGAGCCCGGGCGCACCGTC 3083
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 3084 GTTAGTGAGCGGGAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGGCCATTG 3143
Qy 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 3144 GGTTCGGAAGAGGAACCGAAGCATCTGCTGCTGAAGAGACCACAGAAATGGTGTGCA 3203
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3204 GTCTCCAGTTAACCCACTCCCGACACACACAGAGAGGCCACTCCGTTGAGGAGGTG 3263
Qy 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3264 GAAGGTGGCGTACTGACATAGAAAGAGAGAGAGCGGACTCAAGAGGTCTCTCCAGGCA 3323
Qy 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3324 GTGCGAAGAAAGTGAAGAGGAATCCAGCTGCCCTGGCACCGTGGGCCAGAGATGTG 3383
Qy 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3384 CTTACGCTGTGCAGAGACA-----GAGCGAGAAAGACCCAGAGAG 3425
Qy 1058 AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
Db 3426 CAGGCTGAAGCGTGGGTCTGAAAGAAAGAGACGGATGTAGTGTGAAAGTAGATGCTCAG 3485
Qy 1078 GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1097
Db 3486 GAGGCAAAACTGAGCGCTTTTACACAAAGGGAAGGTGGTGGGCGAGACCACCCAGAAAGC 3545
Qy 1098 LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db 3546 TTTGAAAAGCTCCTCAAGTCAAGAGAGCATAGATCCAGTGAGCTTGTAACACTTGT 3605
Qy 1114 GlnVal-----IleLysLeuGlnGlnLeu---MetGluGlnAlaValAla 1127
Db 3606 CAAGCCGAACCTTAGCTGGGGTAAATCAACAGAGATGGTGTGTAACAGCTATCCCC 3665
Qy 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 3666 CCTGACTCGGTGGAACCCCTACAGACAGTGAAGTGTGTAAGAGCAGCCCGTAGCCGAC 3725
Qy 1148 SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla 1167
Db 3726 TTTGAGCAGCCAGCCACACCCAGAAAGAGAGATTTGTGGAATTCATAGGAGAAATGAG 3785
Qy 1168 ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
Db 3786 GTGCATCTGGTA---CCAGTCAGGGGACACAGAGAGAGGAGGAGGAGGAGGAGGAGGAG 3842
Qy 1188 GluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly 1207

```

Db 3843 AGGCTCCAGCA---CCTTCCAGTTTGTGTTCCAGAGAAACCTAAGAACATCAAG 3899
Qy 1208 ---ArgspValleuGluProthrgIngluLeuThralaalaValaProvalleu 1226
Db 3900 ATGGAAGACACTCTAGAGATACAGATTAAGAGGTGTGAGTGAAGATCTATTCATTCTG 3959
Qy 1227 AlaLysThrgIngluValGluIngluValaLysPLeuAspGlyLysVallys 1246
Db 3960 TCAGAGACTGAGGGAGCTCA-----GAGGCTGACAGTATGCTGATGAGAAACCAAA 4013
Qy 1247 -----GluGluIngluValaPheValHisSerGly-----ProAsnSer 1259
Db 4014 GACGTAACATTTTTCAGAAAGACTTGAGGGGTCTATAGACACAGGCTAACAGTACGTCGG 4073
Qy 1260 GluLysAlaalaAspValThrTyraSerSerGluValaMetGlyValaLagLysGluInglu 1279
Db 4074 GAAAGGTGACACGAGATGCCCTTAAGGTGAAGGACAGAAAGAGCTGAATGTAAAG 4133
Qy 1280 LysGluSerThrgIngluValGlnSer-----LeuSerLeuGluGluGlu 1294
Db 4134 GATGATGCTCTTGAACTGACAGAGTCAAGCTAAGTCTCTCCATCCCGCTGAGAGAGAG 4193
Qy 1295 MetGluThraAspValGluLysGluLysArgGluThraLysProGluIngluValaSerGluInglu 1314
Db 4194 ATGTGATGTTCAAGTCAAGAGGAGAGAAACAGAGACAGACCAACCATGTGATGAAGAG 4253
Qy 1315 Gly---GluGluIngluThraAlaAlaProGluHisGluGlyThrTyrgLysProValleu 1333
Db 4254 AAGCTTGAGCAGAGAAACACTGTTACCGTATCTGAAGAGGTCAAGTGAAGAGCTCCAG 4313
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlySer--- 1352
Db 4314 ACAGTGCAATGTCCCATCATAGATGGGCAAAAGAGTGCAGAGTGTGAAGAGAAACCCCT 4373
Qy 1353 ProSerLeuProAspGluAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4374 CCTCCCTGCTTGAAGTCAAGAGAGGAGCATATGACCAAAATTCAGTTCAGACCTTGAG 4433
Qy 1373 ThrThraValThrgIngluThraLagLalaValaGluLysValIle-----GluThraVala 1390
Db 4434 GCATATTCATCTTAACACAGCGGTGCAGAGAGGAGAAAGGTCTTAGAGAAACCTGCAAC 4493
Qy 1391 HisSerGluThrgLysGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4494 ATTTTAGAAACAGGTGAACGTTGGAGCTGCAGAGTGCACATTAGTTCGAGAGAGAA 4553
Qy 1411 SerSerAlaThrgLysIleHisThraLeuGluHisAlaGluAspThraValProleuGly 1430
Db 4554 TCCTCTGAAAAAATGAAGACTTTCGGCTCATCCAGGGGAAAGATGCTGTGCCACAGGG 4613
Qy 1431 ProGluSerGluAlaGluSerIleProIleIleValIleThraProAlaProGluSerThraLeu 1450
Db 4614 CCCGACTGTCAGGCAAAATCGACACAGATGATGTCGTACTACCAAGAAAGCTTAA 4673
Qy 1451 HisProAspLeuGluIngluLysIleSerAlaSerGluArgGluArgSerGluGluAsp 1470
Db 4674 AGTTCGAGCTGGAAGAGAAACACATCATCTGAAGTGAAGTCAATGATGAAGTCGAT 4733
Qy 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThraAlaIleGluLysValleu 1490
Db 4734 GAGCAGGTGCTTCCAGAGAGTC-----AAAGTAGGTGACCAATTGAGGATTTA--- 4784
Qy 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValleu 1508
Db 4785 -----GAGCCTGAAAAAATGGGATTTTGAACTTGAGACCAAAAGCAGTAACTGTCCAA 4838
Qy 1509 AsnValIleGluThraAlaAspGluPheAlaArgThr---GluThraAlaProGluThra 1527
Db 4839 AACATCATCTCAGACAGCCGTTGACCGATTGTATCGTACGAGAGAAACACCCGCAAAAG 4898
Qy 1528 HisAlaLysAspSerGluThrgIngluValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4899 TTGACGCTGTAGTTACAGACAAAGCTCAGTGATTAAGTGCAGACCGAGCGCTGGA 4958

```

```

Qy 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVal 1562
Db 4959 CAGAAACGAGAAAGAGAGAGAGAGAACCTCAGAGCCTCTGCACAGATGAAGAACCAATT 5018
Qy 1562 LProGluProArgGlu-----AspLeuGlnValleuThraValle 1575
Db 5019 ACTTGAGCCAAAGAGAGACTCAGACTCAACCGCAGTGGACAGACACATTCTGATATTTC 5078
Qy 1575 uGluAlaTrpAlaGluProArgLysCysLeuProArgLeuGluInleuLysAlaProVal 1594
Db 5079 AAAGCATGATGAGAGCTCAGAAAGAACCATGACTGTGAGTAGAGTAAGTTCACCTG 5136

RESULT 3
US-08-994-570-4
; Sequence 4, Application US/08994570
; Patent No. 6090929
; GENERAL INFORMATION:
; APPLICANT: Scott, John D.,
; APPLICANT: Nauer, Brian J.,
; APPLICANT: Klauk, Theresa M.,
; TITLE OF INVENTION: Protein Binding Domains of Gravin
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994,570
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090929and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6605 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 192..5531
; US-08-994-570-4

Alignment Scores:
Pred. No.: 1.51e-288 Length: 6605
Score: 4281.00 Matches: 965
Percent Similarity: 68.6% Conservative: 189
Best Local Similarity: 57.44% Mismatches: 409
Query Match: 53.03% Indels: 117
Gaps: 37

US-09-902-432-4 (1-1596) x US-08-994-570-4 (1-6605)
Qy 1 MetGlyAlaGlySerSerThrgIngluInArgSerProGluIn---ProAlaGlySerAsp 19
Db 192 ATGGGCGCCGGAGGCTCAACCGAGCAGCGACGCCGCGGAGCGCCGCGAGGAGAGCTCC 251

```


Db 2364 GACGGATCTTCTGCTGCCAAGAACTGATCCAGGGCAGGAACTCTCCCGGAG 2423
 Qy ProAlaGlySerProSerGluGluGluGluValSerThrTyrGluSerPheLysArgLeu 752
 Db 2424 CAAGCTGGAAAGCCCTACCGAAGGGAGGGCGTTCCACTCGGAGATCATTTAAAGTTA 2483
 Qy ValThrProArgLysLysSerLysSerLysLeuGluGluValAlaGluAsp----- 769
 Db 2484 GTCAGCGCCAAAGAAAAAATCAAGTCCAGCTGGAAAGAAAAAGCAACATCCATAGCT 2543
 Qy ---SerSerValGluGluLeuSerThrGluLeuGluProSerArgGluGluSerTyrVal 788
 Db 2544 GGGTCTGGTGAACATTCCTCCAGCACTGAACCCGGTAAGAAATAATCTCGGCTC 2603
 Qy SerLysLysPheLeuProGluArgArgLysArgLysArgLysArgLysGluGluGln 808
 Db 2604 TCAATCAAGAAAGTTATCTCTGAGCAAGAAAGAAAGGCCAGATGGGAAACAGAACAA 2663
 Qy AlaThrValGluAspSerGluProValGluLeuSerLysAspProAsnValProAla 828
 Db 2664 GCCCTGTGAAGACGACGAGGCCAACAGGGGCCAACGAAGATGACTCTATGTCCTCGGCC 2723
 Qy ValValProLeuSerGluTyrAsnAlaValGluArgLysMetGlu-----Ala 845
 Db 2724 GTGGTCCCTCTGCTGAGTATGATGCTGTAGAAAGGAGAAATGGAGGCACAGCAAGCC 2783
 Qy GlnGluAsnThrGluLeuProGluLeuGluAlaValTyrValSerGluGluLeuSer 865
 Db 2784 CAAAAGGCGCAGAGACGCCGAGAGAGGCCACCATGAGGTGTCCAAGAGCTCAAC 2843
 Qy LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
 Db 2844 GAGAGTCAGGTTCATATGTGGCAGCAGCTGCTCGTAGAGGGACGAGGGCAGCTACCAT 2903
 Qy ValGluGluArgSerProSerTyrPheSerAlaSerValThrGluProGluGluHisThr 905
 Db 2904 ATTGAAGAAAGGTCTCTCTTGTGATATCTGCTTCACTGACAGAACTTTGAACAAGTA 2963
 Qy AlaGluGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
 Db 2964 GAAGCTGAAGCCGCACTGTAACCTGAGAGGATATTGAAAGAAAGTAATTGACAGAA 3023
 Qy GlnThrProValLeuThrGlnThrLeuProGluGluLysAspAlaHisAspSerMetVal 944
 Db 3024 GAACCCCGCAGGTTCATCACTCTCCAGCAAGAACAGAGAGGCCGCGGCGCAGCGTTC 3083
 Qy ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
 Db 3084 GTTAGTGAAGCCGAATTGACCCCCGAAAGCTGTACAGCTGCAGAAATCTGACGGGCCATTG 3143
 Qy ArgThrGluGluValThrGluAlaSerGluAlaGluGluThrThrAspMetValSerAla 984
 Db 3144 GGTTCCGAAGAAAGAACCCGAAGCATCTGCTGCGTAGAGACCAAGAAATGGTGTACAGA 3203
 Qy ValSerGluLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
 Db 3204 GTCTCCCACTTAACCACTCCCAAGACACACAGAGAGAGGCCCATCCGGTGCAGGGGGTG 3263
 Qy GlnSerGluValLeuAspThrGluGluGluGluGluGluGluGluGluGluGluGluGlu 1024
 Db 3264 GAAGGTGGGTACTGACATAGAAAGAGCAAGAGGGCGGACTCAAGAGGCTCTCCAGGCA 3323
 Qy ValAlaAspLysValLysGluSerGluValProAlaThr----- 1038
 Db 3324 GTGGCAGAAAATGTAAGAAAGAAATCCAGCTCCCTGCGCACCGGTGGCCAGAAATGTG 3383
 Qy ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluGluGlu 1057
 Db 3384 CTTACGCTGTGAGAGAGCA-----GAGCGCAAGAAAGACCAAGAGAG 3425
 Qy AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln 1077
 Db 3426 CAGGCTGAAGCCTGGGTCTGAAGAAAGAGACGGAATGATGTGAAGATGATGCTCAG 3485

Qy GlnAlaGluAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer 1078
 Db 3486 GAGCGCAAAAAGTCCCTTTTACCAAGGGAAGGTGTGGTGGAGACCACCCCGAAGAAC 3545
 Qy LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys 1113
 Db 3546 TTTGAAAAAGCTCTCTCAAGTACAGAGAGCATAGGTCCAGTGAAGCTTTGAACCACTTGT 3605
 Qy GlnVal-----IleLysLeuGlnGluLeu---MetGluGlnAlaValAla 1127
 Db 3606 CAAGCCGAACCTTAGCTGGGGTAAATCAACAGAGATGTGATGAACAGCATATCCCC 3665
 Qy ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
 Db 3666 CTTGACTGGTGGAAACCCCTACAGACAGTGAAGACTGATGGAAACACCCCTGTAGCCAC 3725
 Qy SerAspThrAlaAspGlyThrGlnGluAspGluThrIleAspSerGlnAspSerLysAla 1167
 Db 3726 TTTGACGACACAGGCACAAACCAAGAAACGAGATTGTGAAATTCATGAGAGAAATGAG 3785
 Qy ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu 1187
 Db 3786 GTGCATCTGTA---CAAGTCAAGGGGCACAGAAAGAGAGGACAGTTCCCGACAGAAAGAG 3842
 Qy GlnProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluGluProGly 1207
 Db 3843 AGGCCTCCAGCA---CCTTCCAGTTTGTGTTCAGAGAAACATTAAGAAACATCAAG 3899
 Qy ---ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProValLeu 1226
 Db 3900 ATGGAAGACACTTAGACATACAGATTAAGAGGTGATGAGAAACGTGTATCCATTGTG 3959
 Qy AlaLysThrGluValGlyGlnGluGluGluValAspTyrLeuAspGlyGluLysValLys 1246
 Db 3960 TCAAGAAGTGAAGGGGACCTCA-----GAGGCTGACCAAGTATGCGATGAGAAACCAA 4013
 Qy -----GlnGluGlnGluValPheValHisSerGly-----ProAsnSer 1259
 Db 4014 GACGTACCATTTTTCGAAGGACTTGAGGGGTCTATAGCACAGGCTATAACGTACGTGCG 4073
 Qy GlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGlu 1279
 Db 4074 GAAAGGTTCACACTGAAGTGTCCCTTAAAGTGAAGGACAGAAAGCTGAATGTAAAG 4133
 Qy LysGluSerThrGluValGlnSer-----LeuSerLeuGluGluGluGlu 1294
 Db 4134 GATGATGCTCTTGAACCTGACAGGTCAAGCTCAATCTCCATCCCGGTGAGAGAGAG 4193
 Qy MetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu 1314
 Db 4194 ATGTAGTTCAGGTCAAGTGAAGGAGAGAAACAGAAAGCGCAACCATGTGATTAAGAG 4253
 Qy Gln---GlnGlnGluThrAlaAlaProGluHisGluGluGluThrTyrGlyLysProValLeu 1333
 Db 4254 AAGCTTGAGCAGAAACAGCTGTACCGTATCTGAAGAGTTCAGTACAGAGTCTCCAG 4313
 Qy ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
 Db 4314 ACAGTGAATGTCCCATCATCATATGCGGCAAGAAAGAGTTCAGAGTTTGAAAGAGCCCT 4373
 Qy ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
 Db 4374 CTTCCCTGCTGATGTCAGAGGAGGCAATAGCACCAAAATTTCAAGTTCAGAGCTGTAG 4433
 Qy ThrThrValThrGlnThrAlaGluAlaValAlaGluLysValIle-----GluThrValVal 1390
 Db 4434 GCATCATTCACCTTAACAGCGGTGCAAGAGAGAAAGTCTTAGAAGAACTCCCAAC 4493
 Qy IleSerGluThrGluGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
 Db 4494 ATTTTAGAAACAGTGAACAGTGGAGCGCTGCAAGTGCACATTTAGTTCTTGAAAGAGAA 4553

```
QY 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4554 TCCTCTGAAAAAATGAAGACTTTGCGCGTCATCCAGGGGAAGATGCTGCGCCACAGGG 4613
QY 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
Db 4614 CCGGACTGTGAGCAAAATCGACACCACTGATAGTATCTCTACTACCAAGAAAGGCTTA 4673
QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluAsp 1470
Db 4674 AGTTCCGACCTGGAAGGAGAGAAAAACCATCATCTGAAGTGAAGTGCAGATGAAGTCAT 4733
QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4734 GAGCAGTTCCTGCCAGGAGGTC-----AAAGTGAGTGTAGCAATTTGAGGATTTA-- 4784
QY 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4785 -----GAGCCTGAAATGGGATTTTGGAACTTGAGACCAAAAGCAGTAACCTTGTCAA 4838
QY 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4839 AACATCATCCAGACAGCGGTGACCACTTTGTAGGTACAGAGAAACAGCCACCGAAATG 4898
QY 1528 HisAlaTyraSpSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4899 TTGACGCTCTGAGTTACAGACACAAAGCTCACGTGATAAAGCTGACAGCCAGCGCTGGA 4958
QY 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa 1562
Db 4959 CAGGAACCGAGAAAGAGAGAGGAGGAACTCAGCGCTCTGACAGGATGAACACCAAT 5018
QY 1562 lProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5019 ACTTCAGCCAAAGAGGAGTCAAGTCAACCGCAGTGGGACAGCACACTTCGTATATTC 5078
QY 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5079 AAAGACATGATGTAAGCCTCAGAAAAGACCATGACTGTTGAGGTAGAAGGTTCCACTG 5136

RESULT 4
US-08-973-462-2
; Sequence 2, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5361
; TYPE: DNA
; ORGANISM: P. falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(5361)
US-08-973-462-2

Alignment Scores:
Pred. No.: 8,51e-23 Length: 5361
Score: 462.00 Matches: 323
Percent Similarity: 36.36% Conservative: 293
Best Local Similarity: 19.07% Mismatches: 646
Query Match: 5.72% Indels: 432
DB: 4 Gaps: 68
```

```
US-09-902-432-4 (1-1596) x US-08-973-462-2 (1-5361)
QY 53 GlnLysAsnGlyGlnLeuSerValAsnGlyValAlaGluGlnGlyAspValHisVal 72
Db 283 CAAGTAATGGTGAATAGTAGT-----GAAGAAGTAAGGAAAAAATT 327
QY 73 GlnGluGluAsnGlnGluGlnGlu---GluGluValValAspGluAspValGlyGln 91
Db 328 CTTGACTTATTAGAAGAAGAAATACATTAACTGAAAGTGTAGATGATAATAAAATTTA 387
QY 92 ArgGluSerGluAspValArgGluLys-----AspArgValGluGlu----- 105
Db 388 GAAGAAGCCGAAGATATAAGGAAAAATATCTTTAATAGTAATATAGAAGAACCAAGAA 447
QY 106 -----MetAlaAlaAsnSerThrAlaValGluAspIle 116
Db 448 AATATTATTGCACATTTATTAAATATATTGGACAAATTCAGAAAAACAGAAAGTGTA 507
QY 117 ThrLysAspGlyGln-----GluGluThrSerGluIleIleGluGlnIleProAla 133
Db 508 TCAGAAATATGTACACAGTCAGTGATGAACCTTTTAAATGAATTTATAAATAGTGTAGATGTT 567
QY 134 SerGluAsnAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGly 153
Db 568 AATGGAGAAAGTAAAGAAAAATATTTTGGAG-----GAAAGTCAAGTTAATGACGATATT 621
QY 154 PheLysLysValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGlu 173
Db 622 TTTAATAGTTTAGTAA-----AGTGTTCACACAGAACACACAC 663
QY 174 LysSerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGluAlaGluAla 193
Db 664 AATGTTGAAGAAAAAGTTCAAGAAAGTGTAGAAGAAATGACCAAGAAAGTGTAGAAGAA 723
QY 194 SerVal-----GlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGlu 211
Db 724 AATGTAGAAGAAATGTAGAAGAAATGACGCGGAAGTGTAGCCTCAAGTGTGTGAAGAA 783
QY 212 Ser-----AlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThr 229
Db 784 AGTATAGCTTCAAGTGTGATGAAAGTATAGATTCAAGTATTGAAGAAATGTAGTCTCA 843
QY 230 LeuLysGlnGlu----- 233
Db 844 ACTGTTGAAGAAATCGTAGCTCCAGTGTGTAGAAGTGTGCGCTCCAAGTGTGAAGAA 903
QY 234 -----GlnSerSerThrGluIleProLeuGlnAlaGluSerAspGlnAlaGluGlu 251
Db 904 AGTGTAGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA 963
QY 252 GluAlaLysAspGluGlyGluGluLysGlnGluLysGluProThrLysSerProGluSer 271
Db 964 AATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 1023
QY 272 ProSerSerProValAsnSerGluThrThrSerSerPheLysLysPhePheThrHisGly 291
Db 1024 ATCGTAGCTCCAACCTGTTGAAGAAATCGTAGCTCCAACCTGTTGAAGAAATGTAGTCTCA 1083
QY 292 TrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspLeuGlu--- 310
Db 1084 AGTGTGTAGAAGTGTGCGCTCCAAGTGTGTGAAGAAAGTGTAGAAGAAATGTTGAAGAA 1143
QY 311 ---ThrAlaGluLysArgLysGluGlnGluAlaGluLysValAspGluGluGluLysGlu 329
Db 1144 AGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA 1203
QY 330 LysThrGluProAla-----SerGluGluGlnGluProAlaGluAspThrAspGln 346
Db 1204 AATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 1263
QY 347 AlaArgLeuSerAlaAspTyrGluLysValGluLeuProLysGlnValGlyAsp 366
Db 347 AlaArgLeuSerAlaAspTyrGluLysValGluLeuProLysGlnValGlyAsp 366
```

```
Db 1264 ATC---GTAAGTCCAACTGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATTT----- 1314
QY 367 LeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLys 386
Db 1315 GTAGCTCCAAAGTGTGTAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAAT 1374
QY 387 MetGluAlaHisGluValValAlaGluValHisValSerThrValGluLysThrGlu 406
Db 1375 GTTGA---GAAAGTGTGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGA 1428
QY 407 GluGluGlnGlyGlyGlyGluValGluGluGlyGlyValValGluGluGlyGlyGlu 426
Db 1429 GAAAGT-----GTAAGTGAAGAAATGTTGAAGAAAGTGTGAAAGAA 1455
QY 427 SerLeuProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAla 446
Db 1456 ACTGTA-----GCTGAAATGTT 1473
QY 447 GluGluLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThr 466
Db 1474 GAAGAAAGTGTAGCTGA----- 1491
QY 467 AspLeuSerProGluGluLysThrLeuProLysHisProGluGlyLeValSerGluVal 486
Db 1492 -----AATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTT 1545
QY 487 GluMetLeuSerSerGlnGluLysValGlnGlySerProLeuLysLysLeuPhe 506
Db 1546 GAAGAAATGTGCTCCAACTGTTGAAGAAATGTAGCTCCAACTGTGAAGAAATGTTGA 1605
QY 507 SerSerSerGlyLeuLysLysLeuSerGlyLysGlnGlnGlyLysArgGlyGlyGly 526
Db 1606 GCTCCAAAGTGTGTGAAGAAAGTGTGCT----- 1632
QY 527 GlyAspGluGluProGlyGluLysGlnHisIleHisThrGluSerProGluSerAlaAsp 546
Db 1633 -----CCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCT 1680
QY 547 GluGlnLysGlyGluSerSerAlaSerSerProGluGluPro-----Glu 561
Db 1681 GAAAAATGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGA 1740
QY 562 GluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGlu 581
Db 1741 GAAAGTGTAGCTGAAATGTTGAAGAAATGTAGCTCCAACTGTTGAAGAAATGTAGCT 1800
QY 582 GlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProThrAlaSerPheLys 601
Db 1801 CCAACTGTT-----GAAAGAAATTTAGCTCA-----ACTGTTGTA 1836
QY 602 LysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGluGluGlu 621
Db 1837 GAAAGTGTGCTCCAAAGT-----GTTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGT 1890
QY 622 Leu---GluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
Db 1891 GTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAAT 1950
QY 641 GlnAspGluVal-----LysThrValGlyGluGluGlnLysProGluGluLysProLysArg 658
Db 1951 GTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCA----- 1995
QY 659 ArgValAspThrSerValSerThrProGluAlaLeuIleCysValGlySerSerLysLysArg 678
Db 1996 -----ACTGTTGAAGAAATTTAGCTCCAAAGTGTGTGAAGAAAGTGTG 2037
QY 679 AlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSer 698
Db 2038 GCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGA----- 2085
QY 699 HisArgAlaGluGluLysArgLysAspLysGluAlaGlyThrAspAlaValProAlaSer 718
Db 2086 -----GCTGAAAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGT 2130
QY 719 ThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSer 738
Db 2131 GTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATGTAGCTCAACT 2190
QY 729 GluGluGluGlyValSerThrThrProGluSerPheLysArgLeuValThrPro----- 755
Db 2191 GTTGAAGAAATCGTAGCT-----CCAAGTGTGAAGAAATGTAGCTCCAAAGTGTGTA 2244
QY 756 -----ArgLysLysSerLysSerLysLeuGluGluLysAlaGluAspSerSerVal 772
Db 2245 GAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTAGCT 2304
QY 773 GluGlnLeuSerThrGluLysGluProSerArgGluGluSerThrProValSerLysLys 792
Db 2305 GAAATGTGAAGAAAGTGTGCAAAATGTTGAAGAAAGT----- 2346
QY 793 PheIleProGlyArgArgLysArgAlaAspGlyLysGlnGlnAla-----Thr 810
Db 2347 -----GTAAGTGAAGAAATGTTGAAGAAAGTGTAGCTCCAACT 2382
QY 811 ValGluAspSerGlyProValGluLysGlnLysAspProAsnValProAlaValVal 830
Db 2383 GTTGAAGAAATGTAGCTCCAAAGTGTGAAGAA-----AGTGTAGCT 2424
QY 831 ProLeuSerGluThrArgAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGlu 850
Db 2425 CCAAAGTGTGAAGAAAGTGTGCAAAACGTTGCACAAATTTATCAGCAAACTTT--- 2481
QY 851 LeuProGlnLeuLeuGluAlaValThrValSerGluGluLeuSerLysThrLeuValHis 870
Db 2482 TTTAGTAATTTATTTAGTGTGATC---GAACTGAGGAAATTAAGACAGATATTAAT 2538
QY 871 ThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSer 890
Db 2539 GAGATAGAA-----GAAGTAAAGAAAT--- 2562
QY 891 ProSerThrPheSerAlaSerValThrGluProLeuGlnHisThrAlaGlyGluAlaMet 910
Db 2563 -----GTAGTCACCACTAGTAAGAAAGTGTGAAGAAAGTGTGAAGTGTGA 2613
QY 911 ProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProValLeuThr 930
Db 2614 ACTACTTTTAGTAATATTTAGAGAG-----ATACAGAAATTAATTAATTAAT 2664
QY 931 GlnThrLeuProGluGly---LysAspAlaHisAspAspMetValThrSerGluValAsp 949
Db 2665 GATACATATAGAGAAATTTGAAGAACTCCACCAAGAAATGTATTAAGCCGCTTAGAA 2724
QY 950 PheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluVal 969
Db 2725 AATACCAGTAAGTAAGAGAAAGAAAGAAAGTAATATTAAGTGTAAAT---GAAAGAGTA 2778
QY 970 ThrGluAlaSerGlyAlaGluLysThrThrAspMetValSerAlaValSerGln----- 987
Db 2779 AAGAA-----GAGTCCCTACCACTTAAATGAAGAACTGTGAAGACGACAGAA 2826
QY 988 -----LeuThrAspSerProAspThrThrGluGluAlaThrProVal 1001
Db 2827 GAAAGAGCGCAAAATACATTAATCGAATAATTTGAATAATTTAGAGAA----- 2874
QY 1002 GlnGluValGluSerGlyValLeuAspThrGluGluGluValArgGlnThrGlnAlaIle 1021
Db 2875 AATGCAGTGAAGAAATGAAGAAATGTTGCAGAAATTTAGAGAAATTTAAACCAACTGTA 2934
QY 1022 LeuGlnAlaValAlaAspLysValLysGluGluGluSerGlnValProAlaThrGlnThrVal 1041
Db 2935 TTTAATTAATCTGATTTAGATTAAGTATAGAGAACAGTAAATTAAGCGAGAAAGTTTAGAA 2994
QY 1042 GlnArgThrGlySerLysAla-----LeuGluLysValGluGluValGlu 1056
Db 2995 AACAAATGAATGATTAAGCAATTTTGTAGTAATAATTTGATTAATTTGAAGAAATTAACA 3054
```

Qy	1057	Glubsp	-----SerGluValLeuLeuAspGlu	1066
Db	3055	GAATAATTTATAACAGGTATGTTTCGAAGTATAGAAACCAGTATAGTAATCCAAATCAGAA	3114	
Qy	1066	LysGluLysAspVal	-----Met	1071
Db	3115	GAAGAAGTGAATTTGAATGAAATGCGTGTAGTTCGATTTTAGATAATATAGAAAATATG	3174	
Qy	1072	ProLysGly	-----ProValGlnGluLala	1079
Db	3175	AAAGAAGGTTTATTAATAATATAGAAAATATTTCAAGTACTGAAGGTGTTCAAGAACT	3234	
Qy	1080	GlyAlaGluHisLeuAlaGlnGly	-----	1087
Db	3235	GTAACCTGAACATGTGACAAATAATATATGCGATGTTGATGTTCTCTGCTATCAAAAGAT	3294	
Qy	1088	-----SerGluThrGlyGlnAlaThrProGluSerLeuGluValPro	1101	
Db	3295	CAATTTTACGAATATTAATGAGGCAGCGGGTTGAAGAAATATGTTTTTAATTTGGAA	3354	
Qy	1102	GluValThrAlaAspValAspHisValAlaThrCysGlnValIleLys	-----Leu	1118
Db	3355	GATGTAATTTAAAGTGAAGTGAATTAATCTGTGAAGAAATTAAGGATGAACCGGTT	3414	
Qy	1119	GlnGlnLeuMetGluGln	-----AlaValAlaProGluSerSerGluThrLeuThr	1135
Db	3415	CAAAAAGAGGTAGAAAAGAAACCTGTAGTATTAATGGAAGAAATGGAAGAAATATTGTA	3474	
Qy	1136	AspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln	1155	
Db	3475	GATGTTATGAGGAGAAAGAAAGATTTAAACAGACAAG---ATGATAGATGCGCATGAA	3531	
Qy	1156	GlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGln	1175	
Db	3532	GAATCCATAGAAATATCTTCAGATCTTAAGAAGAAACTGAATCTATTAAAGATAAA	3588	
Qy	1176	ValThrGluGluGluAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsn	1195	
Db	3589	-----CAAAAAGATGTTCCACTAGTTGTTGAAGAA-----	3618	
Qy	1196	ValProAlaGlnGluGluHisGlyGluGluProGlyArgAspValLeuGlu	-----Pro	1213
Db	3619	-----GTTCAACACAAATGATATGGATGAAGCTGTTGAGAAGTTTGTAGAAATTCGAAAAAT	3672	
Qy	1214	ThrGlnGlnGluLeuThrAlaAlaAlaValProVal	-----Leu	1226
Db	3673	ATGGAAGAGAGGTTAATGAAGTCTGTTGAATTAATGACATTAATCTAGCAAACTTATT	3732	
Qy	1227	AlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeu	-----AspGlyGluLysVal	1245
Db	3733	GAAGAAACTCAAGAGTTAAATGAAGTAGAAGCAGATTTAATAAAGATATGGAATAATTA	3792	
Qy	1246	LysGlu	-----GluGlnGluValPheValHisSerGlyPro	1257
Db	3793	AAAGAATTAGAAAAAGCATTTATCAGAAGATTCCTAAAGAATAATATAGATGCCAAAGATGAT	3852	
Qy	1258	AsnSerGlnLysAlaAlaAspValThrTyrAsp	-----SerGluValMet	1272
Db	3853	ACATTAGAAAAAGTTATTGACGAGGNACATGATATAACGACGAGCTTGGATGAAGTTGTA	3912	
Qy	1273	GlyValAlaGlyCysGlnGlu	-----	1279
Db	3913	GAATTTAAAGATGCTCAAGAGAACAAGATCGAAAAAGTATCTGATTTAAAGATCTTGAA	3972	
Qy	1280	-----LysGluSerThrGluValGlnSerLeuSerLeuGluGlu	1292	
Db	3973	GAAGATATATTAAGAAGGTAAAGAAGAACTCAAGAAGCTTGAAGTGAATAATTTAGAAAT	4032	
Qy	1293	GlyGlu	-----MetGluThrAspValGluLysGluLysArgGluThrLysPro	1308
Db	4033	TATAAGAATTAATAAACTATTGAACACAGATATTTTGAAGAGAGAAAAGAAATAGAAAA	4092	
Qy	1309	GluGlnValSerGluGlyGluGlnGluThrAlaAlaProGluHisGluGlyThrTyr	1328	

DECLASS

```

US-08-973-462-1
; Sequence 1, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBESIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0123-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 1
; LENGTH: 6152
; TYPE: DNA
; ORGANISM: P. falciparum
US-08-973-462-1

```

Alignment Scores:

Pred. No.: 1.02e-22 Length: 6152
 Score: 462.00 Matches: 323
 Percent Similarity: 36.36% Conservative: 293
 Best Local Similarity: 19.07% Mismatches: 646
 Query Match: 5.72% Indels: 432
 DB: 4 Gaps: 68

US-09-902-432-4 (1-1596) x US-08-973-462-1 (1-6152)

QY 53 GlnlysaenglgluleuserValaInglyValaInglyaspValaHisVal 72
 Db 527 CAGTAAATGTCGATTAAGT-----GAAAGAGTAAAGGAAAAATT 571
 QY 73 GlnGlnGlnaInglnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 91
 Db 572 CTGACTTATTGAAGAGGAAATACATTACTGAAAGTGCATGATGATATAATAATTAA 631
 QY 92 ArgGlnSerGlnaspValaArgGlnlys-----AspArgValGlnGln----- 105
 Db 632 GAAGAAAGCCGAGATATTAAGAAATAATCTTATTAACTTAATAGAGAACCAAGAAAGAA 691
 QY 106 -----MetAlaAlaSerThrAlaValGlnGlnPile 116
 Db 692 AATATTATTGACAAATTATTAAATAATATTGACAAATTCAGAAAAACAGAAAGTGA 751
 QY 117 ThrlysaSpGlyGln-----GlnGlnThrSerGlnlelleGlnGlnleProala 133
 Db 752 TCAGAAAAATCTCAACTGATGATGAACTTTTAAAGCAATTATTAAATAGTGTAGATT 811
 QY 134 SerGlnAsnAsnValGlnGlnMetValGlnProAlaGlnSerGlnAlaAsnAspValGly 153
 Db 812 AATGCGAAGTAAAGAAATATTTTGGAG-----GAAAGTCACAGTTATGACGATATT 865
 QY 154 PheIysIysValPheIysPheValGlyPheIysPheThrValIysIysAspIysaIngln 173
 Db 866 TTTTAATAGTTTGTGTAATA-----AGTGTTCACAGAAACAAACACAC 907
 QY 174 LysSerAspThrValGlnleuLeuThrValIysIysaSpGlnGlyGlnGlnGlnGlnGln 193
 Db 908 AATGTTGAAAGAAAAAGTTGAAGAAAGTGTGAAGAAATGACACAAAGAAAGTGTGAAGAA 967
 QY 194 SerVal-----GlyAlaGlyAspHisGlnGlnProSerValGlnThrAlaValGlyGln 211
 Db 968 AATGTGAAGAAATATGTAAGAAATGACAGCGAAAGTGTGACCTCAAGTGTGAAGAA 1027
 QY 212 Ser-----AlaSerIysGlnSerGlnleuIysGlnSerThrGlnIysGlnGlnGlnGln 229
 Db 1028 AGTATAGCTTCAAGTGTGATGAAGATATGATTCAGATTTGAAGAAATGTAGCTCCA 1087
 QY 230 LeuIysGlnGln----- 233
 Db 1088 ACTGTTGAAAGAAATCTAGCTCCAAGTGTGTGAAGAAAGTGTGCTCCAAGTGTGAAGAA 1147
 QY 234 -----GlnSerSerThrGlnIysProLeuGlnAlaGlnSerAspGlnAlaGlnGln 251
 Db 1148 AGTGTGAAGAAATATGTAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAA 1207
 QY 252 GlnAlaIysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 271
 Db 1208 AATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 1267
 QY 272 ProSerSerProValAsnSerGlnThrThrSerSerPheIysIysPhePheThrHisGly 291
 Db 1268 ATCGTAGCTCCACTGTGAAGAAATCGTAGCTCCACTGTGAAGAAATGTAGCTCCA 1327
 QY 292 TrpAlaGlyTrpArgIysIysThrSerPheIysIysSerIysGlnAspIysaIngln----- 310
 Db 1328 AGCTGTTGAAGAAAGTGTGCTCCAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAA 1387
 QY 311 ---ThrAlaGlnIysArgIysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 329

Db 1388 AGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA 1447
 QY 330 LysThrGlnProAla-----SerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 346
 Db 1448 AATGTTGAAGAAATGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 1507
 QY 347 AlaArgLeuSerAlaAspTrpGlnIysValGlnleuProLeuGlnAspGlnValGlyAsp 366
 Db 1508 ATC---GTAGCTCCAACTGTTGAAGAAATCTAGCTCCAACTGTTGAAGAAATCT----- 1558
 QY 367 LeuGlnAlaSerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 386
 Db 1559 GTAGCTCCAACTGTTGAAGAAAGTGTAGCTCCAACTGTTGAAGAAAGTGTGAAGAAAT 1618
 QY 387 MetGlnAlaHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 406
 Db 1619 GTTGA-----GAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGA 1672
 QY 407 Gln 426
 Db 1673 GAAAGT-----GTAGCTGAAATGTTGAAGAA 1699
 QY 427 SerLeuProProGlnIysLeuAlaGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 446
 Db 1700 ACTGTA-----GCTGAAATGTT 1717
 QY 447 GlnGlnleuMetLysSerArgGlnMetCysAlaSerGlyGlyAspHisThrGlnleuThr 466
 Db 1718 GAAGAAAGTGTAGCTGAA----- 1735
 QY 467 AspLeuSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 486
 Db 1736 -----AATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTT 1789
 QY 487 GlnMetLeuSerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 506
 Db 1790 GAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTTGA 1849
 QY 507 SerSerSerGlyLeuIysLeuIysLeuSerGlyIysIysGlnGlnGlnGlnGlnGlnGln 526
 Db 1850 GCTCCAACTGTTGTGAAGAAAGTGTGCT----- 1876
 QY 527 GlnAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 546
 Db 1877 -----CCAACTGTTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCT 1924
 QY 547 GlnGlnIysGlyGlnSerSerAlaSerSerProGlnGlnGlnGlnGlnGlnGlnGlnGln 561
 Db 1925 GAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGA 1984
 QY 562 GlnThrThrCysLeuGlnIysGlyProLeuGlnAlaProGlnAspGlyGlnAlaGlnGln 581
 Db 1985 GAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCT 2044
 QY 582 GlyThrThrSerAspGlyGlnIysIysArgGlnGlyIleThrProThrAlaSerPheIys 601
 Db 2045 CCAACTGTT-----GAAAGAAATGTAGCTCCA-----AGTGTGTA 2080
 QY 602 LysMetValThrProIysValArgValArgArgProSerGlnSerAspIysGlnGlnGln 621
 Db 2081 GAAAGTGTGCTCCAACT-----GTTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGT 2134
 QY 622 Leu-----GlnIysValIysSerAlaThrLeuSerThrAspSerThrValSerGlnMet 640
 Db 2135 GTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAAT 2194
 QY 641 GlnAspGlnVal-----LysThrValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 658
 Db 2195 GTTGAAGAAATGTAGCTCCAACTGTTGAAGAAATGTAGCTCCA----- 2239
 QY 659 ArgValAspThrSerValSerTrpGlnAlaLeuIleCysValGlySerSerIysIysArg 678
 Db 2240 -----ACTGTTGAAGAAATGTAGCTCCAACTGTTGTAGAAAGTGTG 2281

Qy	679	AlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyClyAspSer	698
Db	2282	GCTCCAAGTGTGAAGAAGTGTAGAAGAAATGGTTGAAGAAAGTGTA-----	2329
Qy	699	HisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValproAlaSer	718
Db	2330	-----GCTGAAAATGTTGAAGAAAGTAGCTGCAAAATGTTTCAGAAAGT	2374
Qy	719	ThrGlnGluGlnAaspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSer	738
Db	2375	GTAGCTGAAAAATGTTGAAGAAAGTAGCTGAAAAATGTTGAAGAAATCGTAGCTCCAAC	2434
Qy	739	GluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrPro-----	755
Db	2435	GTTCGAAGAAATCGTAGCT-----CCAACTGTTGAAGAAATTTGTAGCTCCAAGTGTGTA	2488
Qy	756	-----ArgLysLysSerLysSerLysLeuGluGluLysAlaGluAaspSerSerVal	772
Db	2489	GAAAGTGTGGCTCCAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCT	2548
Qy	773	GluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLys	792
Db	2549	GAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGT-----	2590
Qy	793	PheIleProGlyArgArgLysLysArgAlaAaspGlyLysGlnGluAla-----Thr	810
Db	2591	-----GTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAAGTGTGCCAAC	2626
Qy	811	ValGluAaspSerGlyProValGluIleAenGluAaspAspProasnValProAlaVal	830
Db	2627	GTTTGAAGAAATGTAGCTCCAAGTGTGAAGAA-----AGTGTAGCT	2668
Qy	831	ProLeuSerGluTyrrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGlu	850
Db	2669	CCAAAGTGTGAAGAAAGTGTGTGTAACACGTTGTCAACAATAATTATCACACAATCTT---	2725
Qy	851	LeuProGlnLeuLeuGlyAlaValTyrrValSerGluGluLeuSerLysThrLeuValHis	870
Db	2726	TTAAGTAATTTATTAGTGTGTATC---GAACTGAGGAAATAAGGCACGATATATTAAT	2782
Qy	871	ThrValSerValAlaValIleAaspGlyThrArgAlaValThrSerValGluGluArgSer	890
Db	2783	GAGATACAA-----GAAGTAAAGAAAAAT---2806	2806
Qy	891	ProSerTrpIleserAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMet	910
Db	2807	-----GTAGTCACCACACTACGAAAAACGTAGAGAAACTACAGCTGAAAGTGTA	2857
Qy	911	ProProValGluGluValThrGluLysAaspIleAlaGluGluThrProValLeuThr	930
Db	2858	ACTACTTTTAGTAACATATTAGAGGAG-----ATACAGAAATACTATTACTTAAT	2908
Qy	931	GlnThrLeuProGluGly---LysAaspAlaHisAaspAspMetValThrSerGluValAasp	949
Db	2909	GATACTATAGGAAAAATTTAGAAGAACTCCACGAAAAATGATTAAAGTCGCTTTAGAA	2968
Qy	950	PheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluVal	969
Db	2969	AATACCCAAAAGTGAAGAGGAAAAAGAAAGTAATAGATGTGAAT-----GAAGAA	3022
Qy	970	ThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaValSerGln-----	987
Db	3023	AAAGAA-----GAGTGTGCTACACTTTAAATAGAAACTGTGGAAACAGCGAGAA	3070
Qy	988	-----LeuThrAaspSerProasppfThrThrGluGluAlaThrProVal	1001
Db	3071	GAAAGAGCGCAATAACAATTACGGAAATATTGTAATAATTTGAAGAA-----	3118
Qy	1002	GlnGluValGluSerGlyValLeuAaspThrGluGluGluArgGlnThrGlnAlaIle	1021
Db	3119	AATCCAGTGAAGATTAATGAAATGTTGCAGAGAAATTTAGAGAAATTAAGCAACTGTGA	3178

Qy	1022	LeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrVal	104
Db	3179	TTTAATACTGTATAGATAAAGTAGAGGAAACAGTAGAAATTTACGGGAGAAAGTTTGA	3238
Qy	1042	GlnArgThrGlySerLysAla	1056
Db	3329	ACCAATCAATCGATAAAGCAATTTTGTAGTGAATATTTGATATGTAAAGGAATACAA	3398
Qy	1057	GluAsp	1065
Db	3299	GAAATTTTAAACAGGTATGTTTGAAGTATAGAAACCAAGTATAGTAATCCATCAGAA	3358
Qy	1066	LysGluLysAspVal	1071
Db	3359	GAAAGGTTGATTTGAATGAAATGTGTTAGTTCGATTTTAGATAATATAGAAAAATAT	3418
Qy	1072	ProLysGly	1079
Db	3419	AAGAAGTTTATTAATAATAATAGAAAATATTTCAAGTACTGAAGGTGTTCAAGAACT	3478
Qy	1080	GlyAlaGluHisLeuAlaGlnGly	1087
Db	3479	GTAACATGACATGTAGCAACAAAATGTATATGCGATGTGGATGTTGATGTTCTGCTATCAAAAGAT	3538
Qy	1088		1101
Db	3539	CAATTTTAGGAATATTAATAGCAGCAGCGGGTTGAAAGAAATGTTTTTAATTTTCGAA	3598
Qy	1102	GluValThrAlaAspValAspHisValAlaThrCysGlnValIleLys	1118
Db	3599	GATGTATTTAAAGTCGAAGTCAATGTAATTACTGTAGAAAGAAATTAAGGATGAACCGGTT	3658
Qy	1119	GlnGlnLeuMetGluGln	1135
Db	3659	CAAAAGAGCGTAGAAAAGCAACTGTTAGTATTTAGGAAGAAATGGAAGAAATATTGTA	3718
Qy	1136	AspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln	1155
Db	3719	GATGTATTAGAGGAAGAAAAGAAAGATTTAAACAGACAAAG---ATCATAGATGCAGTAGAA	3775
Qy	1156	GlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValAlaGlnSerGln	1175
Db	3776	GAATCCATAGAAATATCTTCAGATCTTAAGAGAAACCTGAATCTTATTAAAGATATAA	3832
Qy	1176	ValThrGluGlnGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsn	1195
Db	3833	-----GAAAAGATGTTTCACCTAGTTGTTGAAGAA-----	3862
Qy	1196	ValProAlaGlnGluHisGlyGluGluProGlyArgAspValLeuGlu-----Pro	1213
Db	3863	-----GTTCAAGACAATCATATGATGGAAGTGTGAGAAAGTTTGAATTTGAAAAAT	3916
Qy	1214	ThrGlnGlnGluThrAlaAlaValProVal	1226
Db	3917	ATGGAAGAGGAGTAAATGAAGATGCTGTTGAATATAAATATGATCATCTAGCAAACTTATT	3976
Qy	1227	AlaLysThrGluValGlyGlnGlyGluValAspTrpLeu---AspGlyGluLysVal	1245
Db	3977	GAAGAACTCAAGAGTTAAATGAAGTAGACGAGATTTAATAAAGATATGGAANAATTA	4036
Qy	1246	LysGlu-----GluGlnGluValPheValHisSerGlyPro	1257
Db	4037	AAAGAAATAGAAAAAGCATTATCAGAAGATTTAAAGAAATAATAGATGCAAAAGATGAT	4096
Qy	1258	AsnSerGlnLysAlaAlaAspValThrTyrAsp-----SerGluValMet	1272
Db	4097	ACATTAGAAAAAGTTATTGAAGAGGAACATGATATAACGACGACGTTGGATGAAGTTGTA	4156
Qy	1273	GlyValAlaGlyCysGlnGlu-----	1279
Db	4157	GAATTTAAAGATGTCGAAGAGACGAAGATCGAAAAGTATCTGATTTAAAGATCTTGAA	4216
Qy	1280	-----LysGluSerThrGluValGlnSerLeuSerLeuGluGlu	1292

```

Db 4217 GAAGATATATTAAAGAACTAAAGAAATCAAGAACTGAAAGTGAATTTAGAAAGT 4276
Qy 1293 Glylu-----MetGluThrAspValGluYsGluYsArgGluThrYsPro 1308
Db 4277 TTTAAAGAAATTTAAACCTTTGAAACAGATATTTTGAAGAGAAAAAGAAATAGAAAAA 4336
Qy 1309 GluGlnValSerGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1328
Db 4337 GATCATTTTGAAGAAATTCGAGAGAGAGCTGAAGAAATTTAAAGATCTTGAACAGATATA 4396
Qy 1329 GlyYsProValLeuThrLeuAspMetProSerSerGlu----- 1341
Db 4397 TTTAAAGAAAGTATCTTCATTAGAGTTGAGAAAGAAAAAAATTAGAGAGATACGAA 4456
Qy 1341 ----- 1341
Db 4457 TTTAAAGAGAGTAGAACATTTAATAGTGGTATGCCATATATAAGGTTTGAGAGAA 4516
Qy 1342 -----ArgGlyYsAlaLeuGlySerLeuGlyGlySer 1352
Db 4517 GATGATTTAGAAAGAGTAGATGATTTTAAAGAGAGTATTTAGACATGTTTAAAGGAGAT 4576
Qy 1353 ProSerLeuProAspGlnAspYsAlaGlyCySIlleGluValGlnValGlnSerLeuAsp 1372
Db 4577 ATGGAATTGAGGATATGATTAAG-----GAAAGTTTAGAA 4612
Qy 1373 ThrThrValThrGlnThrAlaGluAlaValGluYsValIlleGluThrValIlleSer 1392
Db 4613 GATGTACACACAAACTTGAGAGAAAGAGTTGATTC---TTAAAGATGTTTATCTACT 4669
Qy 1393 GluThrGlyGluSerProGluCySValGlyAlaHis----- 1404
Db 4670 GCATTGCGCATGATGAGAAACAAATGAAACAAAGAAAAAAAGCTCAAGACCTTAAGTTG 4729
Qy 1405 -----LeuLeu-----ProAlaGluYsSerSerAlaThr 1414
Db 4730 GAAGAAGTATTTTAAAGAGAGGTTTAAAGAAAGAAACAAAGAAAAAAATTAACAAAAAAG 4789
Qy 1415 GlyGlyHisThrThrLeuGlnHisAlaGlu-----AspThrValProLeuGlyProGlu 1432
Db 4790 AAGTAAAGGTTTATTTAAGATTAAGAAACCAAAAGATGAATA---GTAGAAAGTTGAA 4846
Qy 1433 SerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeuHisPro 1452
Db 4847 ATCAAGATGAGAGATATA-----GAAAGAGATGAGAGAA 4882
Qy 1453 AspLeuGlnGlyGluIleSerAlaSerGlnArgGluYsSerGluGluGluGluGluGlu 1472
Db 4883 GATATGAAAGAAAGATATTAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 4942
Qy 1473 AspAlaGlyProAspAlaAspGlyYsGluSerThrAlaIlleGluYsValLeuYsAla 1492
Db 4943 GATATAGTGAGACAAAGAT-----GAAAGTATAGATTTAATAGTCCAAAAA 4990
Qy 1493 GluProGluIleLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1506
Db 4991 GAGAAACGCATTTGAAAGGTTTAAAGCAAAAAAGAAAAAATTAA 5032

```

```

STREET: Edelhertweg 15
CITY: PH Leijstad
STATE:
COUNTRY: The Netherlands
ZIP: NL-8219
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS v.6.0
SOFTWARE: WordPerfect v. 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/119,125A
FILING DATE: 20-SEP-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL92/00054
FILING DATE: 19-MAR-1992
APPLICATION NUMBER: NL 9100510
FILING DATE: 21-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Handal, Anthony H.
REGISTRATION NUMBER: 26275
REFERENCE/DOCKET NUMBER: SMITHHE119125
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 838-8589
TELEFAX: (203) 838-8794
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6744 base pairs
TYPE: Nucleic acid with corresponding amino acids
STRANDEDNESS: single stranded
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORGANISM: Streptococcus suis type II (pathogenic)
FEATURE:
OTHER INFORMATION: Extracellular factor related protein (EF*) gene
FEATURE:
NAME/KEY: promoter -35 region
LOCATION: bp 66 to 71
FEATURE:
NAME/KEY: promoter -10 region
LOCATION: bp 89 to 94
FEATURE:
NAME/KEY: promoter -35 region
LOCATION: bp 153 to 158
FEATURE:
NAME/KEY: promoter -10 region
LOCATION: bp 176 to 181
FEATURE:
NAME/KEY: ribosome binding site
LOCATION: bp 350 to 356
FEATURE:
NAME/KEY: signal peptide
LOCATION: bp 361 to 498
FEATURE:
NAME/KEY: start of repetitive units R1-R11
LOCATION: bp 2869, 3097, 3292, 3520, 4087, 4381, 4609, 4837, 5065, 5293, 5521:
FEATURE:
NAME/KEY: start of repetitive Asn-Pro-Asn-Leu sequences
LOCATION: bp 2932, 3160, 3355, 3583, 4150, 4444, 4672, 4900, 5128, 5356, 5584:
FEATURE:
NAME/KEY: dyad symmetry regions
LOCATION: from bp 6554 to 6566 and from bp 6571 to 6583
FEATURE:
NAME/KEY: dyad symmetry regions
LOCATION: from bp 6611 to 6625 and from bp 6631 to 6644
US-08-119-125A-2
Alignment Scores: 7.62e-20 Length: 6744
Pred. No.:

```


Score:	421.50	Matches:	333
Percent Similarity:	36.25%	Conservative:	268
Best Local Similarity:	20.08%	Mismatches:	664
Query Match:	5.22%	Indels:	393
DB:	1	Gaps:	71

US-09-902-432-4 (1-1596) x US-08-119-125A-2 (1-6744)

QY	4	GlySerSerThrGluGlnArgSerProGluGlnProAlaGly-----	17
Db	2056	GGTACAAACACTAAGATTACTTCAGAACTCGATCCTACGGGAGCAACTGCAGAAACCAAGGA	2115
QY	18	-----SerAspThrProSerGluLeuValLeuSerGlyHisGlyProAlaAla	33
Db	2116	GATGACGGTCAATCTTCAACTAAGTTTAACGTTTAAGATTACGACCGACCTGCTTACA	2175
QY	34	GluAlaSerGlyAlaAlaGlyAspProAlaAspAlaAspProAlaThrLysLeuProGln	53
Db	2176	GAAGGTACCGGCATTAT-----	2193
QY	54	LysAsnGlyGlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGln	73
Db	2194	-----AAGCTTCGTGTTTGA	2208
QY	74	GluGluAsnGlnGlu---GlyGlnGluGluGluValValAsp-----	88
Db	2209	GAAGATAAAGTATCTCTTTGTCGACAGGGGAACTCTGTTGATGGAATAAACCAGAAAT	2268
QY	89	ValGlyGlnArgGluSerGluAspValArgGluLysAspArgValGluGluMetAlaAla	108
Db	2269	GTAGTTTGACATCTGTAAAGATTACCTTCGTAAA-----CATGCTACGGTGCTCA	2319
QY	109	AsnSerThrAlaValGluAspIleThrLysAspGlyGlnGluGluThrSerGluIleIle	128
Db	2320	ACACCACTTCTGTTGAAATCCAGCTAACTTAACGCCAGAGAAAGAAAGCCGAGTTATT	2379
QY	129	GluGlnIleProAlaSerGluAsnAsnValGluGluMetValGlnProAlaGluSerGln	148
Db	2380	GCTCAAAATCAAGAAAGAACACGACGACCAACAGATTTGAAGGGCTTGCACAGATTCA	2436
QY	149	AlaAsnAspValGlyPheLysLysValPheLysPheValGlyPheLysPheThrValLys	168
Db	2437	-----GCATTTCAGTTTAAAC	2451
QY	169	LysAspLysAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGly	188
Db	2452	TCAGATGGT-----ACTGTGTCAGTTGACTACAGT	2481
QY	189	GluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGluThrAla	208
Db	2482	GCCGGTGGTGTCAATGTTGATGGTGCACACACATTAATAAGAAATGCTACCAAACTTG	2541
QY	209	ValGly-----GluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGln	226
Db	2542	GCAGATACACGGATGAAGCAAGCAAGAAATCGACAAATAATGCTGAACATAAAAA	2601
QY	227	GluGlyThrLeuLysGlnGluGlnSer-----SerThrGluIleProLeuGln	242
Db	2602	GCTATCGAAGCAAAACGGGATGAAGCGTTTCTAAAATTGATGATGACATTCCTTGAGA	2661
QY	243	AlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLysGlnGlu	262
Db	2662	GCAGAACAGACAGAGCGTGTCAAGGATCGCGTTGCTGCAGCTGCTGGGGATGCTTTGAA	2721
QY	263	LysGluProThrLysSerProGluSerProSerProValAsnSerGluThrThrSer	282
Db	2722	GAATTAGAACACAGGCGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	2781
QY	283	SerPheLysLysPhePheThrHisGlyTyrAlaGlyTyrArgLys-----	297
Db	2782	GAATCAATGATGCTAAGACTAATGGTGCAGATTAACTGACAGCTGCAGAGCAGTAGGC	2841
QY	298	LysThrSerPheLysLysSerLysGluAspLeuGluThrAlaGlu---LysArgLys	316

Db	2842	GAATAAGCTATTAAACAGCGAAGGAAAAAGAACTGGCAAAAGCAGAAAGTTGAAAAA	2901
QY	317	GluGlnGluAla---GluLysValAsp-----GluGluGlu-----	327
Db	2902	GCATTCGAGGCATTGGAAAAAGTTTACAATAATCAAACTTGTGTAAGAGAGAGAAAA	2961
QY	328	-----LysGluLysThrGluProAlaSerGluGluGlnGluProAla	341
Db	2962	GCATACTTTTGATGATATTAAAGAACTTAAGAAAGTTGCAAGTTGAGAAAAATCAATAATGCT	3021
QY	342	GluAspThr-----AspGlnAlaArgLeuSerAlaAspTyrGlu	354
Db	3022	GAATAACTGCTGAAATTACGAGCAATTTGACAAAGCGGAAATTCATACATGAAGAT	3081
QY	355	LysValGluLeuProLeuGluAspGlnValGlyAspLeuAlaSerSerGluGluLys	374
Db	3082	GTTATTATACGACGCCCACTTGATGCTTTGTAATAAGCTTTGAAAAAGATAGCGAAGAACT	3141
QY	375	CysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGlnGluValVal	394
Db	3142	AAGGCAGCTATT-----GATGCTAATCCAACTTAACT	3174
QY	395	AlaGluValHisValSerThrValGluLysThrGluGluGlnGlnGlyGlyGlyGlu	414
Db	3175	CCGGAAGAGAAAGCGAAGCTATTGCTAAGGTAGTAGAGCTTGT-----AATAAT	3225
QY	415	AlaGluGlyGlyValValValGluGlyThrGlyGluSerLeuProGluLysLeuAla	434
Db	3226	GCTGATCTGACATTTTGTGGAAGCTCCCCAGAAACAGTT-----	3267
QY	435	GluProGlnGluValProGlnGluAlaGluProAlaGluGluMetLysSerArgGlu	454
Db	3268	-----CAAGCAGTGGAGGATAAGGCTGAC-----AAAGATCTTGCCAAAGTAGAACTT	3315
QY	455	MetCysValSerGlyGlyAspHisThrGlnLeu-----ThrAspLeuSerProGlu	471
Db	3316	CAAGCAGCAGCAGCAGCGTGGGAGAAAGCATTTGAAGCGCTCAACTCCAACCGAAGTT	3375
QY	472	GluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSer	491
Db	3376	GAGAAAGATGATAGCTAAGAAG-----GCAGTA	3402
QY	492	GlnGluArgIleLysValGlnGlySerProLeuLysLys-----	504
Db	3403	GAAGACGCGGTTAAGGTGGCGACAGACGCTATTGATAAGCGCTCAACTCCAACCGAAGTT	3462
QY	505	---LeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLysArg	523
Db	3463	GACACAGCAGCAGCAGCGATGGAGTGAAGGCTATTGATGCGAAGAGAGTTTAAAGCTACTCAG	3522
QY	524	GlyGlyGlyGlyAsp-----GluGluProGlyGluTyrGlnHisIleHisThrGlu	540
Db	3523	AAGATGCTAAGAACAAAGATTGCCAAAGAGCAGAAATCAGCTAAGAAAGGATTGACGAC	3582
QY	541	SerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluPro	560
Db	3583	AATCCAAACTTGACTCCAGATGAGAAG---GAATCAGCTAAGAAATCGAGTGAAGAGCGC	3639
QY	561	GluGlu-----ThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGlu	578
Db	3640	GCTAAGGTAGCAACAGCCGCTATTGATAAGCA-----TCAACTCCAGATGCAAGTTCA	3693
QY	579	AlaGluGluGlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTyrAla	598
Db	3694	GTAGAAGAG-----GACAAAGGTAGCAGCTATCAATTTGATTACTGCCAAGGCA	3744
QY	599	SerPheLysLysMetValThrProLysLysArgValArgProSerGluSerAspLys	618
Db	3745	GATGCTAAAGGTGCTATTGCTGCTAAGTTGGCAGATGAAATCAAGAGCTCGAAGATAAG	3804
QY	619	GluGluGluLeuLys---ValLysSerAlaThrLeuSerSerThrAspSerThrVal	637

Db 3805 CAAGCAGAAAGCAAAAGCTATTCGATGGTCAACTATGACTAATGAGGAAAGCAATC 3864
Qy 638 SerGlu-----MetGlnAspGluValLysThrValGlyGluGluGluLysProGlu--- 654
Db 3865 GCTAAGAAAGGCTCTTCAGATGTTGTA-----GATTAAGAAAGCAAGAGCTT 3912
Qy 655 GluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySer 674
Db 3913 GAAGACGACGCTTACGACCAACAAATGAGATTTCATGAAAGCT-----ACT 3957
Qy 675 SerLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeu 694
Db 3958 ACTACAGAAAGAAAGCGAAGCGGGAACCTTGCTGGGGAAGAGCTTGACAGACACA--- 4014
Qy 695 GlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAla 714
Db 4015 GGTAAAGAGCTAGAGATCAGTTGAAATTGGCTAAGATTAAGAAATTACCTAAGAGACA 4074
Qy 715 ValProAlaSerThrGlnGluGluAspGlnAlaGlnGlySerSerSerProGluProAla 734
Db 4075 ATCCGAAAGCAAGAAAGCAAGCACTTAAATATGATGAGAAACTTCAGAGATACGCC 4134
Qy 735 GlySerProSerGluGlyGluGlyValSerThrTrpGlu-----SerPheLysArg 751
Db 4135 AAAGCTATGAGAGACAAATCCAAACTTGTCAAGATGAGATTAAGCAAGCGAAATTAAAG 4194
Qy 752 LeuValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAspSerSer 771
Db 4195 CTAACTGACGCTGTGGCAAAACTTTCAGAACCATTCGTGACATCAATGAGATGACCTAG 4254
Qy 772 ValGlu-----GlnLeuSerThrGlnIleGluProSerArgGluGluSerTrp 787
Db 4255 CAGAAAGCAAGAAAGCTCAAGCCCTAGCAAGATCTTGAAAGAAAGCTTAAGAAACACAGAA 4314
Qy 788 ValSerIleLysLysPheIleProGlyArgArgLysArgAlaAspGlyLysGlnGlu 807
Db 4315 ATTGCGATTAAGAGCTCCGATTTAGTTGACTATCTTGTAAGATGTGTGAGCTTGA 4374
Qy 808 GlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValPro 827
Db 4375 ---GCTACTAAACAGAT----- 4389
Qy 828 AlaValAlaProLeuSerGluTrpAsnAlaValGluArgGluLysMetGluAlaGlnGly 847
Db 4390 GCTAAGAACAAAGATTCTTAAGATGCAAGCGCTGTPAAAGCAATTCGACAGCAATCCA 4449
Qy 848 AsnThrGluLeuProGlnLeuGluAlaValIleValSerGluGluLeuSerLysThr 867
Db 4450 AAC-----TTGACAGACCGACAGAAAGAAC 4476
Qy 868 LeuValHisThrValSerValAlaValIle-----AspGlyThrArgAlaValThr 884
Db 4477 TTCACCGATGCGGTAGATGACAGAAATAGCCAAAGTAAAGCAAGCAATTCAGCTGCAAC 4536
Qy 885 Ser-----ValGluGluLysSerProSerTrpIleSerIleSerAlaThrGluPro 901
Db 4537 AGCCGACGAGATTCTTAAAGAAAGAGAGATGCGGTGTGCACCAATTCGACAGAAAGTGT 4596
Qy 902 LeuGlnHisThrAlaGlyGluAlaMetProProVal-----GluGluValThrGluLys 919
Db 4597 CTTCGACGACGCTTAAACAAAGATTAAGATTAAGATTGCTTAAAGATGCAAGCGCTGTA 4656
Qy 920 AspIleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAla 939
Db 4657 GAAGCAATTGGC---TCCATCCAACTTGACAGACGCA-----GAGAAGAAAGAAC 4704
Qy 940 HisAspAspMetValThrSerGluValAspPheThrSerGluAlaValThr---AlaThr 958
Db 4705 TTCACCGATGCGGTAGATGACAGAAATAGCCAAAGCTTAAGCAAGCAATTCAGCTGCAAC 4764
Qy 959 GluThrSerGluAlaLeuArgThrGlnGluValThrGluAlaSerGlyAlaGluGluThr 978
Db 4765 AGCCGACGAGATTCTTAAAGAAAGAGATGACAGGTGTTGCACCAATTCGACAGAAAGATGT 4824

Qy 979 ThrAspMetVal-----SerAlaValSerGlnLeuThrAspSerProAsp 993
Db 4825 CTTCGCGAGCTTAAACAAAGATGCTAAGATTAAGATGCTTAAAGAAATCGCAGCTGCTAAG 4884
Qy 994 ThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGlu 1013
Db 4885 TCAGGCATTGACGCGCAATCCA-----ACTTGACAGATGACAGAGAG 4926
Qy 1014 Glu-----GluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGlu 1031
Db 4927 GATTCAGCTAAGAAAGCAATTGATGCTAAGCTTCGACAGATGATGATGATGCT 4986
Qy 1032 GluSerGlnValProAlaThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLys 1051
Db 4987 TCACACATCCAGTCGAAAGCGCAATCGGCAAGACAGCAAGAGC----- 5028
Qy 1052 ValGluGluValGluGluAspSerGluValLeuAlaSerGluLysGluLysAspValMet 1071
Db 5029 GTAGGTCAATCGCCCAAGATCTTTCAGCCAGCGAAACAGATGCTTAAGACAAAGATT 5088
Qy 1072 ProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGly 1091
Db 5089 -----GCCAAAGAGTTGCCGCGACGTAAGAAAGCAATT 5121
Qy 1092 GlnAlaThrProGlu-----SerLeuGluValProGluValThrAlaAspVal 1107
Db 5122 GATGCCAATCCGAACTTATCAGATGACAGAGAAAGAGCTTTCAGAAACCGTGAATCA 5181
Qy 1108 AspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnAlaValAla 1127
Db 5182 GATGCTAAGCTACACAGATGCAATT-----GATGCTTCAACAAAGT 5223
Qy 1128 ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp 1147
Db 5224 CCAGTCGAAGCGCAATCG-----GCAGAGGCAAGAGCGGTGATTCATCGCC--- 5271
Qy 1148 SerAspThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSerGlnAsp 1164
Db 5272 CAAGATGTTCTTGACGACGCAAGCAAGATCTTAAGATTAAGATTGCTTAAGAAATCCAG 5331
Qy 1165 SerLysAlaThrAlaAlaValAlaArgGlnSerGlnValThrGluGluGluAlaAlaThrAla 1184
Db 5332 GCTGCTAAGTACGCCATTGACCGCAATCCAAACTTGACAGATGACAGAAAGAACTCACT 5391
Qy 1185 GlnLysGlu-----GluProSerThrLeu 1192
Db 5392 AAGAAAGCGGTAGATGACAGATGCTTAAAGCTGCGACAGATGCAATGATGATTCACAAAGT 5451
Qy 1193 ProAsnAsnValProAlaGlnGluGluHisGlyGluGluPro---GlyArgAspValLeu 1211
Db 5452 CCAGTCGAAGCGCAATCGGCAAGAGCAAGAGCGGTAGCGCCATGCCAAAGACATTCCTT 5511
Qy 1212 GluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAlaLysThrGluVal 1231
Db 5512 GATGCCGCGCAACAAAGAT-----GCTAAGAACAAAGATT 5544
Qy 1233 GlyGlnGlnGlyGluValAspTrpLeuAspGlyLysValLysGluGlnGluVal 1251
Db 5545 GCTTAAAGAGCA-----GAATTCGCTAAGTCA 5571
Qy 1252 PheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTyraAspSerGlu--- 1270
Db 5572 GTCAATTGACTCCAAATCCGAATCTTACAGATGCTTAAGAGAGCGGCTTAAATTAATT 5631
Qy 1271 -----ValMetGlyValAlaGlyCysGlnGlnGluLys 1280
Db 5632 GATTAAGCTGTTGAGAGAGCAATTTTATATCAATGCTGTTAGAACTTATCAAGAGTTG 5691
Qy 1281 GluSerThrGluValGlnSerLeuSerLeu----- 1290
Db 5692 GAAAAATCAAACTTCATGCGACGCTTAATTAAACAGCTGCGAAAGTAAACACACAGTG 5751

QY 1291 -----GluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysPro 1308
Db 5752 GTTGATCAAAATACTGACTGAAAAAATAATCTCTGATCAAGGCATTCCTTAAAGAG 5811
QY 1309 GluGlnValSerGluGluGlyGluGlnGluThrAlaAlaProGluHisGluGlyThrTyr 1328
Db 5812 AACATAACCTCCATAAGGACAGACGATTAAATGTTTCTAAAGATGCTTCAGTGACAAAT 5871
QY 1329 GlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySer 1348
Db 5872 AAATATCCAGATGAACTATTGATTGCTATCACA----- 5907
QY 1349 LeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnVal 1368
Db 5908 -----GTAGAAGTT---GTG 5919
QY 1369 GlnSerLeuAspThrThrValThrGlnThrAlaGlu----- 1380
Db 5920 AAGCAGCAGATAAACTGCTCCTACGGTCGCAATGATGCAAGGTAATATTGTGATT 5979
QY 1381 ----AlaValGluLysValIleGluThrValVal-----IleSerGluThrGlyGluSer 1397
Db 5980 GTACCGCTGAAAAAGCTGTGAGCTTGTGTTTCATACGTAGATAACAATGGTAAGTCG 6039
QY 1398 ProGluCysValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHis 1417
Db 6040 CAACTGTAGTT-----GTTACGAAAGGTACCGATGGTTA 6075
QY 1418 TrpThrLeuGlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSer 1437
Db 6076 TGGACA-----GCAAGTAATACAGTG----- 6096
QY 1438 IleProIleIleValThrPro-----AlaProGluSerThrLeuHis 1451
Db 6097 -----GTGATTGGACCCCTGCTGACTGGGCAAGTAATCGTTCAGGTTCGTATTAAAG 6150
QY 1452 ProAspLeuGlnGlyGluIle-----SerAlaSerGlnArgGluArgSerGluGluGlu 1469
Db 6151 CCA-----GGTACAGTTGTTACAGCATCTCTAAAGACGAGTTGGAAATAGTTCT 6201
QY 1470 AspLysProAspAlaGlyProAspAla---AspGlyLysGluSerThrAlaIleGluLys 1488
Db 6202 GATTCACGACAGCTGAAGTTGAGTAGAGTAGACGAAATAATCTCGACAGGAGTGAAA 6261
QY 1489 ValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAsnLysIle 1506
Db 6262 GTTAAATCAGTTACTACAAATGCTAATAATGTTGAGAGAAAGTAAGCAATTA 6315

RESULT 7
US-08-466-390-3
Sequence 3, Application US/08466390
Patent No. 5686562
GENERAL INFORMATION:
APPLICANT: TOUKATLY, GARY
APPLICANT: LIGARD, GRAHAM P
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,390
FILING DATE: 06-JUN-1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6306 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6306
PUBLICATION INFORMATION:
AUTHORS: COMPTON, DUANE A
AUTHORS: SZILAK, ILYA
AUTHORS: CLEVELAND, DON W
TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
TITLE: SEGREGATION OF PROTEINS AT MITOSIS
JOURNAL: J. Cell Biol.
VOLUME: 116
PAGES: 1395-1408
DATE: 1992
US-08-466-390-3

Alignment Scores:
Pred. No.: 1,98e-19 Length: 6306
Score: 415.00 Matches: 373
Percent Similarity: 35.28% Conservative: 255
Best Local Similarity: 20.96% Mismatches: 663
Query Match: 5.14% Indels: 493
DB: 72 Gaps: 72

US-09-902-432-4 (1-1596) x US-08-466-390-3 (1-6306)

QY 28 GlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAspAlaAspPro 47
Db 1050 GGCACCTAGGAGTGGCTAG 1109
QY 48 Ala-ThrLys-----LeuProGlnLysAsn-----GlyGlnLeuSerVa 61
Db 1110 GCAGGACAGAAATGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1169
QY 61 l-----AsnGlyValAlaGluGlnGlyAspVal----- 70
Db 1170 GGAAGAACACTTGTCCAGCTGCAGGATTAACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1229
QY 71 -----HisValGlnGluGlnGlnGlnGlnGln----- 79
Db 1230 TGATGTCTTCAGCTGGAAACCTTGAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1289
QY 80 -----GlnGluGluGluValValAspGluAspValGlnArgGluSerGluAs 96
Db 1290 ACAGCTCCAAAGCCAGGAGTAGAGATGCTGGAGACTGAGCGAGGAGAGAGAGAGAGAGAG 1349
QY 96 pValArgGluLysAspArgValGluGlu----- 105
Db 1350 GCTTGTGAGCGGGCCACTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1409
QY 106 -----MetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyG 121
Db 1410 CTGACAGAGCTCCATCTCCAAACCTCAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1469
QY 121 nGluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValGluGluMe 141
Db 1470 GGCTCATGGGGCCGGTTGACTGCCAGAGGTGGCCCTCTTGACCTCTGAGCTCACCACACT 1529

Oy	141	cValgInProAlaGluSerGlnAlaSerValGlyPheUlySerValPheUlyPheVa	161
Db	1530	CAATGCCACCATCCAGCAACGATCAAGAAAGG-----GC	1565
Oy	161	IGlPheUlySerPheThrValUlySlyAspUlySerAsnGluUlySerAspThrValGlnUle	181
Db	1566	TGGCCTGAAGACAGACAGGCCCAAAAGAAAGCAGGCCCACTAGACACAGACCTTCAA-----	1620
Oy	181	uThrValUlySlyAspGluUlyGlnGluValaGluValaSerValGlyAlaGlyAspHisG	201
Db	1621	-----CAGCAAGAACAGGCGCTCCAGGGCGCTC-----CGCCACCA	1655
Oy	201	nGluProSerValGlnUthrAlaValaGlyUlySerAlaSerUlySglUlySerUlyUlyG	221
Db	1656	G-----GTGAGAGAGCTTAAGACAGTACAGTACCTGAAGCAAGAAAGAGACAGCTTAAGA	1706
Oy	221	nSerThrGluUlySglGlnGluUlyThrUleUlySglnGlnGlnUlySerSerThrGluUlePro	241
Db	1707	GGTACCGGAGAAAGACAGAGGCACT-----AGCCAGACCACTGCCACCAACTGGCCAC	1760
Oy	241	uGlnAlaGluSerAspGlnAlaAla-----GlnUlnGluAlaUlyAspGluUlyGlnG	259
Db	1761	TGCTGCAGAGAGAGGAGAGAGGCTCTTAAGAGAGCGGATGCGGCTCTTAACAGACTGGA	1820
Oy	259	uUlySglGlnUlySglUProThrUlySerProGlu-----SerProSerSerProValAs	277
Db	1821	GGCACTGGAGAGAGAGAAAGGCTGCAAGCTGGAGATTCTGCAGACCAACTCAAGTGGC	1880
Oy	277	nSerGluUthrThrSerSerPheUlySlyPhePheThrHisGlyTPrAlaGlyTPrArgUly	297
Db	1861	TAAATGAACCCCGGACAGCTGCGCCAGCTTACATGACACAG-----GCCACAGGGA	1931
Oy	297	sUlyThrSerPheUlySlySerUlySglUAsp-----AspUlnGluUThrAlaGluUly	314
Db	1932	GAAGCAGAGAGCTGAAGCGGAAGTGAGAGAACTCCAGCGCTGGTTGAACAGCCCGCA	1991
Oy	314	sArgUlySglUlnGluValaGluUlySerValaSerGlnGlnGlu-----UlySglUlyGTh	331
Db	1992	GGAAACGACATGAGGCCAGGCGCCAGGTTCGACAGCTTGAAGCTTCGCGCTGCACCA	2051
Oy	331	rGluProAlaSerGlnGlnGlnGlnUProAlaGlnUAspThrAspGlnAlaArgUleUSeAl	351
Db	2052	GCAAAAGCAACTGAGAAAGAAAGAGGTGGCCCAAGAAAGACAG-----CTCCAGGA	2105
Oy	351	aAspUlyGluUlySerValaGluUleProUlnGluAspGlnValaGlyAspUleGluUlaSe	371
Db	2106	GCAGCTCCAGGCGCTCAAAAGAGCTCTTGAAGGTCACCAAGGGGAGGAGCTTGAA-----	2157
Oy	371	rGlnUlyUlyCysAlaProUlnAlaThrGluValaPheAspGluUlySerMetGluAlaHisG	391
Db	2158	GAGGAGAAAG-----CGCAGGGCTGCAGATGCCCTGGAAAGACAGCAAGCTTGATATC	2210
Oy	391	nGluValValaAlaGluVala-----HisValSerThrValGlnUlyGTh	405
Db	2211	TGAGCTGAAGCAAGACCGGAGGCTGTGTGACACGCAATAGCGGAAAGAAAGAGACT	2270
Oy	405	rGlnUlnGlnGlnGlnGlyGlyGlyGlnUlnGluUlnGlyUlyValaGlnUlyGThG	425
Db	2271	GGAAAGAAAGAGGGCTGGGCGCAAGGGGCTGCAGGCTCATTACTGACGCTTGGGGAAGC	2330
Oy	425	yGluUleUleUProGlu-----UlyUleUlnGluUProGlnGluValaProG	441
Db	2331	CCATCAGGCTGAGACTGAATCTCGCGCGGAGGTCGAGAGGCCATCGGCTGCCACCA	2390
Oy	441	nGluUlnAlaUly-----ProAlaGlnUlnUleUlySerThrArgUlnUleCysValaSerG	460
Db	2391	CACACCTGAGGTGAGTGTGACGCTGTCCAAAGAAATGATGCTGCTGGCGTGCAGGGGA	2450
Oy	460	yAspHisThrUlnUleUThrAspUleUSeProGlnGluUlyUlyThrUleUProHisUProG	480
Db	2451	TGAGGATAGCCAG-----CAAGAGAGAGC-ACAGTATGGCGGCAATGTTCC	2494
Oy	480	uGlyUlnUlyUlySerGlnUlnGluUleUleUSeSerGlnUlnUlyUlyUlyUlnGlnUlySe	500

Db	2495	AGGAAACAGCTGATGACTTTGAA	-----	2516
Qy	500	rProLeuLysLysLeuPheSerSerSerLysLeuLysLysLeuSerGlyLysLysGlnLys		520
Db	2517	-----	-----GGAGGAATGTGGA	2530
Qy	520	EGLYLysArgGlyGlyGlyLysAspGlnGluProGlyGlyTyrGlnHisSerLeuThrGly		540
Db	2531	AGGGCCCCCAGAGAGCTGCAGAGAGCAAGAGAGGTGCAGAGCTAGAAATCCCAACGG		2590
Qy	540	userProGlyLysAlaAspGlnGlnLysGlyLysSerSerSerAlaSerProGlnGluPro		560
Db	2551	AGCTTCA-----GATAGCCGGCAGACAGAACAACTACTAGCTTCATGCGCAAC		2611
Qy	560	oGluGluThrThrCysLeuGlnLysGlyProLeuGlnLysAlaProGlnAspGlyGlnAla--		579
Db	2642	TGGCCAGAGCACT-----CCAGCAGCTCCAAAGAGAGAGTCAAGGCCCAAGAGCTTG		2695
Qy	580	-----GluGlnGlyLysThrThrSerAspGlyLysLysLys		591
Db	2696	CAGATGACCTTCACCTGCAGAGAAATAGTGCTGCCACCGCAAGAGGTGGCCGCT		2755
Qy	591	rgGlnGlyLysLeuProThrAlaSerPheLysLysMetAlaThrProLysLysArgVal		611
Db	2756	TGGAGACCTTGCTCCGAGAGCGAGTGAGACAGAGAGAAACAGCTCCGGAGATTAGTCA		2815
Qy	611	rgArgProSerGluSer---AspLysGlnGlnGlnLysLysSerAlaThr		629
Db	2816	AGGAGCCTCGAGAGGCGAGAGACAGACACGCCAGTGCTGTGAAAGCAACAGGAGCC		2875
Qy	629	hrLeuSerSerThrAspSerThrValSerGluMetGlnAspGlnValLysThrValGly		649
Db	2876	AGTTCTGCAGACACAGCGAGCGCTGCAGAGCTATGAGCGGAGGACAGACAGATGGGCA		2935
Qy	649	LgluLys-----GlnLysProGlnGlnGln		655
Db	2936	ATGAGCTGGAAAGCGCTCCGGCCCGCTGATGAGAGACCGAGGCGACGACGAGGAGCG		2995
Qy	655	-----	-----	655
Db	2996	GMGGCGACAGAGAAAGGAGAGTGGCCGGCGGTACCCAGAGAGGGGGCGGTGCCAGCGTG		3055
Qy	656	-----P	-----	656
Db	3056	ACCTTGCCTCGAGAGAGCGCGCAGAGAGCTTGAGATCGCGCTGCAGAACGCCCTCA		3115
Qy	656	roLysArgArgValAspThrSerValSerThrGlnAlaLeuLecyValGlySerSerL		676
Db	3116	ACGAGCAGCGTGTGAGTTCGTACCTCGACAGAGGACATGCTCATGCTGACGGAAA		3175
Qy	676	YSLysArgAlaArgLysAlaSerSerSerAspAspGlnLysGlyProArgThrLeuGlyG		696
Db	3176	AGGAA-----GCGAAGAGCCAGAGATT--GGCAAGCTTGCTGCTGCTG		3216
Qy	696	LysPheSer---HisArgAlaGlnGlnAlaSerLysAspLysGlnAla-----GlyT		712
Db	3217	GAGGACGCCAGATPAAAGAGCTGAGAGAACTTCGGCAAAACGTGAAGCACTGAAGGAA		3276
Qy	712	hrAspAlaValProAla--SerThrGlnGlnGlnAspGlnAlaGlnGlySerSerSerPro		731
Db	3277	CA-----GCTGGCTTAAGAAAGAAAGAGACACGACTGTGGCTCAGAGACCCAA		3324
Qy	732	GluProAlaGlySerProSerGlnGlnGlyGlnLysValSerThrTrpGlyLysPheLysArg		751
Db	3325	TCTAGAGCTGTGCAGAGACAGACCAACAGGCCCAACTG---GAAACATCGCGGCA		3381
Qy	752	LeuValThrProArgLysLysSerLysSerLysLeuGlnGlnLysAlaGln-----		768
Db	3382	GAGGTGACAGCTGTAACAGCAATGCCAGAAAGCACAGACAGGCTTACAGCTTGA		3441
Qy	769	-----AspSerSerValGlnGlnLeu		775

Db 3442 CGCAGCCTCGAGGCTGAGCGGCTCCCGGCTGAGCGGCACAGTGTCTCTGGAGACTCTG 3501
Qy 776 SerThrGluLeuGluProSerArgGluGluSerTrpValSerIleLysLysPheIlePro 795
Db 3502 CAGGCGCAGTTAGAGAGAGAGCCAGGCTAGAGGACACAGTACAGTGCCTTAGCCTCG 3561
Qy 796 GlyArgGlyLysArgAla-----AspGlyLysGlnGlnAlaThrValGluAsp 813
Db 3562 GCCAAGCGGAGTTGGTCTTCCGACCAAGATCAAGACACACAGCAAGGCTGAAGAT 3621
Qy 814 SerGlyProValGluLeu-----AsnGluAspAspProAsnValProAlaVal 829
Db 3622 GAGTGGAGCCCGAGTGGCGCGCGGCGGCGGAGAGCTGAGAGAAAATAGCCTCATC 3681
Qy 830 ValProLeu---SerGluTyArgAlaValGluArgGluLysMetGluAlaGlnGlyAsn 848
Db 3682 AGCAGCTTGGAGGAGGAGTGTCCATCTGAATCGCCAGGCTCTGGAGAGGAGGG--- 3738
Qy 849 ThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeu 868
Db 3739 ---GAGAGCAAGGAGTTGAAGCGGCTGGTATGCCCGAGTCAGAGAAAGCCAGAAGCTG 3795
Qy 869 ValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGlu 888
Db 3796 GAGGAGAGCTGGCGCTCTCG-----AGGCAG 3822
Qy 889 ArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGlu 908
Db 3823 AGACAGCCAGCA----- 3834
Qy 909 AlaMetProValGluGluValThrGluLysAspIleIleAlaGluGluThrProVal 928
Db 3835 ---ACAGTGCAGAGCTGCAGAACGCAGCTCTGTCTGCGGAGG--- 3876
Qy 929 LeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGlu--- 947
Db 3877 ---AGGTGCAGAGCTCGGGAGGAGCTGAGAAACAGCGGTTGGCTTCAGAAAC 3930
Qy 948 ---ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
Db 3931 CTGCGGCAGGAGCTGACCTCACAGGCTGAGCGTGCAGGAGGCTGGGCCAAGAATTGAAG 3990
Qy 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
Db 3991 GCGTGGCAGGAGAAGTTCTTCCAGAAAGAGCAGCGCCCTCTCCACCTGCAGCTC--- 4044
Qy 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003
Db 4045 -----GAGCACACACGACACAGGCC 4065
Qy 1004 ValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGln 1023
Db 4066 CTGCTGAGTGTGCTGCCAGCTAAGCACCTCTGCCAGCAGCTGCGAGCGCGAGGCC 4125
Qy 1024 AlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
Db 4126 GCTGCCAGAGAAACCCACCTGAGGAGCTGGAG-----CAGAGCAAGCAGGCC 4173
Qy 1044 ThrGlySerLysAlaLeuGlu----- 1050
Db 4174 GCTGGGGGAGCTGGGGCAGAGCTGCTCGGGGCCAGCGGAGGCTTGGGGAGCTGATTCTCT 4233
Qy 1051 -----LysValGluGluValGluGluAspSerGluValLeuAlaSerGluLys--- 1066
Db 4234 CTGCGGCAGAGAGTTGGCAGAGAGGAGGAACAGCTCAGCAGGCTGCGGGCAGAGAGGCC 4293
Qy 1066 ----- 1066
Db 4294 AGCTATGACAGCAGCTGAGCATGCTGAAGAGGCGCATGCGCTGTCTGGCAGAGAGAAC 4353
Qy 1066 ----- 1066
Db 4354 CGGGGGCTGGGTGAGCGGGCCAACTTTGGCCGGCAGTTTCTGGAAAGTGGAGTTGGACCAG 4413

Qy 1067 -----GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeu 1084
Db 4414 GCCCGGAAAGATGATCTCCAAGAGTTGGCAGCGCTACGTCTGCTGAGACCCGCTG 4473
Qy 1085 AlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThr 1104
Db 4474 GCTGAGGTGCGAGGAGACAGACAGACACTCCCGGAGCTGGAGGTG-----ATGACT 4527
Qy 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGln 1124
Db 4528 GCCAAGTATGAGGTGCC-----AAGGTCAAGGTCTCTGGAGGAGG 4569
Qy 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
Db 4570 CAGCGGTTCAGGAAGAGAGGAGCAAACTCACTGCCAGGTGGAAGAACTGAGTAAGAAA 4629
Qy 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSer 1162
Db 4630 CTGGTGACTCTCAACAGCCCAAGGTGACGACAGCAAGCTGAAGCTGTCCAGGCT 4689
Qy 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
Db 4690 CAGGAGGCGAGAGCCAGCAGGAGGCCCGCGCTTCAGGCCCGAGTGAATGAATGCA 4749
Qy 1181 AlaAlaThrAlaGlnLysGluGluProSer-----ThrLeu 1192
Db 4750 GCCCAGTTGAGCCAGAAGGAGCAGGAGCTGAGCACTATAAGCTGCAGATGGAGAAAGCC 4809
Qy 1193 ProAsnAsnValProAlaGlnGluGluHisGlyGluGlu-----ProGly 1207
Db 4810 AAACACATTTATGATCCCAAGAGCAGCAGAACCAAGAGCTGCAGGAGCAGCTGCGGAGC 4869
Qy 1208 ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaAlaValProValLeuAla 1227
Db 4870 CTGAGCAGCTGCAGAGGAAACAAAGAGCTGCGAGCTGAAGCT----- 4914
Qy 1228 LysThrGluValGlyGlnGluGlyValAspTrpLeuAspGlyGluLysValLysGlu 1247
Db 4915 ---GAACGGCTGGGCCATGAGCTACAG-----CAGGCTGGCTGAAGACCAAGGAG 4962
Qy 1248 GluGlnGluValPheValHis-----SerGlyProAsnSerGlnLysAlaAlaAspVal 1265
Db 4963 GCTGAACAGACCTCGCGCCACCTTACTGCCAGGTGCGCAGCTGGAGGCAGAGTTGCC 5022
Qy 1266 ThrTyAspSerGluValMetGlyValAlaGlyCysGln-----GluLysGlu 1281
Db 5023 CATGCAGACAGCAGCTTCGAGACCTGGGCAATTCAGGTGCAACTGATGCTTTAAAG 5082
Qy 1282 SerThrGluValGlnSer-----LeuSer 1289
Db 5083 AGCGTGAGCCCGCAGGCTAAGCCCGAGCTTGAGTATTGACAGCCTGGATCTCGAGC 5142
Qy 1290 LeuGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGlu 1309
Db 5143 TGCAGAGGAGGAGCCCACTCAGCTATCACAGCAAGCTGCTCGT---ACCCAGCCAGAC 5199
Qy 1310 GlnValSerGluGlyGluGlnGlu-----ThrAlaAlaProGluHisGluGly 1326
Db 5200 GGACACAGCTCCTCGAGAACCCAGCTCACCTATCTCCCA-GCGCTGCCCCCAAGGT 5258
Qy 1327 ThrTyGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeu 1346
Db 5259 AGAATCCCTGGAGAGTCTCTACTTACTCCCATCCCTGCTCGGAGTCAAGGCCCCCT--- 5315
Qy 1347 GlySerLeuGlyGlySerProSerLeuPro----- 1356
Db 5316 -----GGAGAGCAGCTGAGCTCCCTGGGAGAGCTTCTCTGAGTCTGGGTCTGTA 5366
Qy 1357 -----AspGlnAspLys 1360
Db 5367 GACCCGCTCGCTCGTGGCGCACCGCAGATCATCAACATCACCATGACCAAGAA--- 5423

Oy	1361	AlAGlGysrIleGIuValGlnValGlnIserLeu-AspThrThrValThrThrAlaG1	1360
		:::	:::
Db	5424	GCTTGAATGTGGAAGAAGCCACAGACGCCCACTATGCTTCTACAGACCGGGTCTGCCTCC	5483
Oy	1380	uAlaValGIuLysValIle-----GIuThrValValIleSerGIuThrGIuGI	1386
		:::	:::
Db	5484	TGCTTCCAGAGCTGTGACCTGGACGACCTCTCTACTCACTCTTAAGCTCGCCCTGGGT--	5541
Oy	1396	uSerProGIuLysValAlaIyAlaHisLeuLeu-----	1406
Db	5542	-TCTCCGATTAATGGAACATCGACGCTGTCTCAAGCTTGCCTGGCTACCGCCCAACATTCG	5600
Oy	1407	----ProAlaGIuLysSerSerAlaThrGIuGIuHisTrrThrLeuGlnHisAlaGIuAs	1425
		:::	:::
Db	5601	CAGTTCGTGCTGTGCTGTCCAGGCCGG-GGTGTCCAGTGGGGGCCCTCCAGAGAAAGAAC	5659
Oy	1425	pThrValPro-----LeuGIuProGIuSe	1433
	:::		
Db	5660	GCTTCTTCAATGAGGACCTTCCAGAGATGAGCTGAGACGTGATGAGAACCCCATCTG	5719
Oy	1433	GIAlaGIuLysTrrLeuProIleIleValThrProAlaProGIuSerThrLeuHisProAs	1453
		:::	:::
Db	5720	CAGAGCTGCAGACGACGCAGATCGATGTGGCCCCCACATCTGAAGAGCTGTATCCCTCGG	5779
Oy	1453	pLeuGIuGIuLysTrrSerAlaSerGIuArgIuArg-----SerGIuGIuAspLysPr	1472
		:::	:::
Db	5780	AGTCCAGGCTTCCCTCTGAGCTGTGGGACCATCAACATGAGAGAGAAACCTGAGAAC	5839
Oy	1472	oAsp-----AlaGIuProArgAlaAspGIu	1480
Db	5840	CCCAAGAGACCTTCGGCCCGAGACCGACAGCATGACGCAATCAAGATAGCCGAGGG	5891

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

? NAME/KEY: CDS
? LOCATION: 1..6306
? PUBLICATION INFORMATION:
? AUTHORS: COMPTON, DUANE A
? AUTHORS: SZILAK, ILYA
? AUTHORS: CLEVELAND, DON W
? TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
? TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
? TITLE: SEGREGATION OF PROTEINS AT MITOSIS
? JOURNAL: J. Cell Biol.
? VOLUME: 116
? PAGES: 1395-1408
? DATE: 1992
US-08-470-950-3

Alignment Scores:	
Pred. No.:	1,98e-19
Score:	415.08
Percent Similarity:	35.28%
Best Local Similarity:	20.96%
Query Match:	5.14%
DB:	1
	Gaps:
	72
	.
	Indels:
	493
	Mismatches:
	665
	Conservative:
	255
	Matches:
	370
	Length:
	6306

US-09-902-432-4 (1-1596) X US-08-470-950-3 (1-6306,

[illegible]

QY 221 nSerThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluLeuProLe 241
Db 1707 GTAGCGGAGACAGAGCAACT-----AGCAGGACCATGCGCCAGCAACTGGCCAC 1760
QY 241 uGlnAlaGluSerAspGlnAlaAala-----GluGluGluAlaLysAspGluGluG1 259
Db 1761 TGCTGACAGGAGCGAGAGGCTCTTAAGGGAGCGGATGCGGCTCTCAAGCAGCTGGA 1820
QY 259 uLysGlnGluLysGluProThrLysSerProGlu-----SerProSerSerProValAs 277
Db 1821 GGCACTGGAGAGGAGAGGCTGCCAAGCTGGAGATTCTGCAGCAGCAACTTCAGGTGGC 1880
QY 277 nSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLy 297
Db 1881 TAATGAAGCCCGGACAGTCCCGCAGACCTCAGTGACACAG-----GCCAGCGGGA 1931
QY 297 sLysThrSerPheLysLysSerLysGluAsp-----AspLeuGluThrAlaGluLy 314
Db 1932 GAAGGCAGAGCTGAGCCGGAAGGTGGAGAACTCCAGGCCGTGTGTGAGACAGCCGCCA 1991
QY 314 sArgLysGluGlnGluAlaGluLysValAspGluGluGlu-----LysGluLysTh 331
Db 1992 GGAACAGCATGAGCCCGCAGCCAGGTTGACAGACTAGAGTTGCAGCTGCGGCTGAGCA 2051
QY 331 rGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAl 351
Db 2052 GCAAAAGCAACTGAGAAAGGAGGTGGCCCGCAGGAGGACAG-----CTCCAGGA 2105
QY 351 aAspTyrgLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSe 371
Db 2106 GCAGCTCCAGGCCCTCAAGAGTCTTGAAGTCCACCAAGGGCAGCCTTGA----- 2157
QY 371 rGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisG1 391
Db 2158 -GAGGAGAAG-----CGCAGGCTGCAGATGCCCTGGAAGACGACGAGGTTGTATCTC 2210
QY 391 nGluValValAlaGluVal-----HisValSerThrValGluLysTh 405
Db 2211 TGAGCTGAAGCGAGACACCGAAGCCTGTGGAGCAGCATAAAGCGGAGCAAGGAGCT 2270
QY 405 rGluGluGluGlnGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrG1 425
Db 2271 GGAAGAGAGAGGCTGGCGCGCAAGGGCTGGAGGCTGATTAAGTCTGAGCTTGGGAGGC 2330
QY 425 yGluSerLeuProProGlu-----LysLeuAlaGluProGlnGluValProG1 441
Db 2331 CCATCAGGCTGAGACTGAAGTCTCGCGCGGAGCTGGCAGAGGCATGCTGCCCCAGCA 2390
QY 441 nGluAlaGlu-----ProAlaGluGluLeuMetLysSerArgLuwetCysValSerGlyG1 460
Db 2391 CACAGCTGAGAGTGAGTGTGAGCAGCTCGTCAAAGAAGTAGCTGCTGGCGGTGACGGTA 2450
QY 460 yAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProG1 480
Db 2451 TGAGGATAGCCAG-----CAAGAGGAGGC-ACAGTATGGCGCCATGTTC 2494
QY 480 uGlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySe 500
Db 2495 AGGAACAGCTGATGACTTTGAA----- 2516
QY 500 rProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLy 520
Db 2517 -----GGAGAAATGTGAGA 2530
QY 520 sGlyLysArgGlyGlyGlyAspGluGluProGlyGlyTyrGlnHisIleHisThrG1 540
Db 2531 AGCCCGCCAGGAGCTGCAGAGGCAAGAGAGAGGTTGGCAGGCATAGATCCCAAGCGG 2590
QY 540 uSerProGluSerAlaAspGluGlnLysGlyGlyGluSerSerAlaSerSerProGluGluPr 560
Db 2591 AGCTCCA-----GATAAGCCGCGCAGCAGAACAACTAGCTAGCTCCATGCCCAACC 2641
QY 560 oGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAla-- 579

Db 2642 TGCCAGAGCACT-----CCAGCAGGTCCAAGAGAAGGAAGTCAAGGCCCAAGAGCTTG 2695
QY 580 -----GluGluGly-ThrThrSerAspGlyGlyLysLysLys 591
Db 2696 CAGATGACCTCTCCACTCTGCAGGAAAAGATGGTCCACCAGCAAGAGAGGTGGCCCGCT 2755
QY 591 rgGluGlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysLysArgValA 611
Db 2756 TGGAGACCTTGTGTCGCAAGCAGGTGAGCAGCAGGAAAACAGCCTCCCGGAGTTAGTCA 2815
QY 611 rgArgProSerGluSer---AspLysGluGluGlu---LeuGluLysValLysSerAlaT 629
Db 2816 AGAGCCTCGAGGCGGAGGAGCAGACCCGAGTGGCTTGAAGAGCAACAGGAGCGCC 2875
QY 629 hrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyG 649
Db 2876 AGTTCTGCAGCACACAGCAGCAGCGCTGCAGGCTATGGAGCGGAGGAGCAGATGGCA 2935
QY 649 luGlu-----GlnLysProGluGlu- 655
Db 2936 ATGAGCTGGAACGGCTGCGGGCGCGCTGTATGGAGAGCCAGGGGCAGCAGCAGGAGGC 2995
QY 655 ----- 655
Db 2996 GTGGCAGCAGAAAGGAGGTGGCGCGCTGACCCAGGAGCGGCGCTGCCAGGCTG 3055
QY 656 -----P 656
Db 3056 ACCTTCCCTGGAAAGCGGCCAGACAGAGCTTGAGATGGCGGTGCAGAACGCCCTCA 3115
QY 656 roLysArgArgValAspThrSerValSerTrpAlaLeuIleCysValGlySerSerL 676
Db 3116 ACAGCAGCGTGTGGAGTTTCGTACCTGCAAGAGGCACTGGCTCATGCCCTCAGCGAAA 3175
QY 676 yLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyG 696
Db 3176 AGAA-----GGCAAGGACCAAGAGTT-GGCCAAGCTTCGTGTCTG 3216
QY 696 lyAspSer-----HisArgAlaGluAlaSerLysAspLysGluAla-----GlyT 712
Db 3217 GAGCAGCCCATAAAGAGCTGGAGGAACCTCGCAACCCGTGAAGCACTGAAGGAA 3276
QY 712 hrAspAlaValProAla-SerThrGlnGlnGlnAspGlnAlaGlnGlySerSerPro 731
Db 3277 CA-----GCTGGCTAAAGAAAAGAGGACACGATCTGCTCAGGAGCCCA 3324
QY 732 GluProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArg 751
Db 3325 TCTGAGGCTGCTGGCAGGACAGAGCCCAAGGCCCAAGCTG---GAAGCACTGCGGGCA 3381
QY 752 LeuValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGlu----- 768
Db 3382 GAGGTGAGCAAGCTGGAACAGCAATGCCAGAAAGCAGCAGCAGCAGCTCAGACCTGGAA 3441
QY 769 -----AspSerSerValGluGlnLeu 775
Db 3442 CGCAGCCTCGAGGCTGAGCGGGCTCCCGGGCTGAGCGGACAGTGTCTCTGGAGACTCTG 3501
QY 776 SerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIlePro 795
Db 3502 CAGGGCCAGTTAGAGGAGAGGCGCCAGGAGCTAGGACAGTGCCTTAGCCTCG 3561
QY 796 GlyArgArgLysLysArgAla-----AspGlyLysGlnGluGlnAlaThrValGluAsp 813
Db 3562 GCCCAAGCGGAGTTGGCTGCTTCCGCCAACCAAGGTATACAGACACACAGCAGGCTGAAGAT 3621
QY 814 SerGlyProValGluIle-----AsnGluAspAspProAsnValProAlaVal 829
Db 3622 GAGTGAAGGCCAGGTGTGCTCCCGCGGCGGCAAGAGGCTGAGAGGAAAAATAGCCTCATC 3681
QY 830 ValProLeu---SerGlyTyraAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsn 848

Db 3682 AGCAGCTTGGAGAGAGGTGTCATCTGAATCGCCAGCTCTCGAGAGAGAGGG--- 3738
 QY ThrGluLeuProGlnLeuLeuGlnValTyrValSerGluGluLeuSerIleu 868
 Db 3739 ---GAGAGCAAGAGTGTAAAGGGGCTGTGATGGCCGATCAGAGAAAGAGCAGAAAGCTG 3795
 QY ValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGlu 888
 Db 3796 GAGGAGAGCTGGCCGTGCTG---AGGCAG 3822
 QY ArgSerProSerTriPleSerAlaSerValThrGluProLeuGlnHisThrAlaGlyGlu 908
 Db 3823 AGACAGCCAGCA--- 3834
 QY 909 AlamelProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProVal 928
 Db 3835 ---ACAGTGCAGAGCTGCAGAAAGCAGAGCTGTCTGTGGCGGAGG--- 3876
 QY 929 LeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGlu--- 947
 Db 3877 ---AGTGCAGAGCTCCCGGAGAGGCTGAGAAACAGCGGTGGCTTCAGAGAAC 9330
 QY 948 ---ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
 Db 3931 CTGCGGCAAGAGCTGCTCAGAGCTGAGCGTGGAGAGAGCTGGGCCAAGATTGAAG 3990
 QY 966 Thr---GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
 Db 3991 GCGTGCAGAGAAAGTTCTCCAGAAAGAGCAGCGCTCCAGCTCGAGAGCTC--- 4044
 QY 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003
 Db 4045 ---GAGCAACACCGACACAGCAGCGCC 4065
 QY 1004 ValGluSerGlyValLeuAspThrGlnGluGluArgGlnThrGlnAlaIleLeuGln 1023
 Db 4066 CTGTGATGATGAGCTGCTGCGACCTAAGACACTTGCAGAGAGCTGAGCGCCAGCAGCGCC 4125
 QY 1024 AlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
 Db 4126 GCTGCGGAGAAACGACACCGTGAAGAGCTGAG---CAGAGCAAGCAGCGCC 4173
 QY 1044 ThrGlySerLysAlaLeuGlu--- 1050
 Db 4174 GCTGGGGGAGCTGGGCGAGAGCTGCGGCGCCAGCGGAGCTTGGGAGCTGATTCT 4233
 QY 1051 ---LysValGluGluValGluLysAspSerGluValLeuAlaSerGluLys--- 1066
 Db 4234 CTGCGGCAAGAGGTGGCAGAGCAGAGCGAAACAGCTCAGCGCGGCGAGAAAGGCC 4293
 QY 1066 --- 1066
 Db 4294 AGCTATGCAAGAGAGCTGAGCATGTGAAGAGCGCATGCTGTGGCAGAGAGAAC 4353
 QY 1066 --- 1066
 Db 4354 CGGGGCTGGGTGAGCGGCGCACTTGGCGGCACTTCTGGAAGTGAAGTGAACAG 4413
 QY 1067 ---GlnLysAspValMetProLysGlyProValGlnGluAlaGlnHisIleu 1084
 Db 4414 GCCCGGAAAGATGATGTCCAAGAGTTGGCAGCCGTACGTCTGATGTGAGACCCGCTG 4473
 QY 1085 AlaGlnGlySerGluThrGlnAlaThrProGluSerLeuGluValProGluValThr 1104
 Db 4474 GCTGAGGTGAGGAGAGAGCAGAGCAGACCTGCGGAGCTGAGAGTG---ATGACT 4527
 QY 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGln 1124
 Db 4528 GCCAAGTATGAGGTGCC---AAGGTCAAAGTCTCGAGAGAGAG 4569
 QY 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
 Db 4570 CAGCGGTTCCAGAGAGAGCAGAAATCTAGCTGCGCCAGGTGAGAAACTGAGTAAAGAAA 4629

QY 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAsp---GluThrIleAspSer 1162
 Db 4630 CTGGCTGACTGTGACCAAGCAGAGAGGTGACAGCAGAGCTGAAGCTGTGACAGCT 4689
 QY 1163 Gln---AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
 Db 4690 CAGGAGGCGAGAGCCAGCAGAGAGGCCAGCGCTTCCAGGCCCCAGCTGAATGAAGTGA 4749
 QY 1181 AlaAlaThrAlaGlnLysGluGluProSer---ThrLeu 1192
 Db 4750 GCCCAGTTGAGCAGAGAGAGCAGAGCAGCTGAGCACTTAAGCTGACATGAGAGAAAGCC 4809
 QY 1193 ProAsnAsnValProAlaGlnGluGlnHisGlyGluGlu---ProGly 1207
 Db 4810 AAACACATTTATGATCCAAAGAGCAGAGCAACCAAGAGCTGCAGAGCAGCTCGGAGC 4869
 QY 1208 ArgAspValLeuGluProThrGlnGlnLeuThrAlaAlaValProValLeuAla 1227
 Db 4870 CTGGAGCAGCTGCAGAGAGAAACAAAGAGCTGCGAGCTGAAGCT--- 4914
 QY 1228 LysThrGluValGlyGlnGluGlyValAlaAspTrpLeuAspGlyGluLysValLysGlu 1247
 Db 4915 ---GAACGCTGGGCGCATGAGCTACG---CAGGCTGGGCTGAAGCAGAGAG 4962
 QY 1248 GlnGlnGluValPheValHis---SerGlyProAsnSerGlnLysAlaAlaAspVal 1265
 Db 4963 GCTGAACAGACCTGCGGCACCTTACTGTCAGAGTCCGACGCTGAGGACAGAGTTGCC 5022
 QY 1266 ThrTyrAspSerGluValMetGlyValAlaGlyCysGln---GluLysGlu 1281
 Db 5023 CATGCAGACACAGACCTTGCAGACCTGGGCAATTCCAGGTGGCACTGATGCTTTAAG 5082
 QY 1282 SerThrGluValGlnSer---LeuSer 1289
 Db 5083 AGCCGTGAGCCCGCAGCGTAAAGCCAGCTGAGCTTGAAGATTGACAGCTGTGATGAGC 5142
 QY 1290 LeuGlnGluGlyGluMetGluThrAspValGlnLysGlyLysGluThrLysProGlu 1309
 Db 5143 TCCGAGAGGGGAGCCCACTCAGTACACAGCAAGCTGCGCTCG---ACCCAGCCAGAG 5199
 QY 1310 GlnValSerGluGluGluGlnGluGln---ThrAlaAlaProGlnHisGluGly 1326
 Db 5200 GGCACCAAGGTCCCTGAGAAACACAGCTTCACTTAATCTCCA-GGCCCTGCCCCCAAGGT 5258
 QY 1327 ThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeu 1346
 Db 5259 AAGATCCCGAGAGTCTTACTTCACTCCATCCATCCCTGCTGGATGACAGGCCCT--- 5315
 QY 1347 GlySerLeuGlyLysProSerLeuPro--- 1356
 Db 5316 ---GGAGAACAGCTGAGACTCCCTGGAGAGCGTCTTCTGAGCTCGGAGTGGAA 5366
 QY 1357 ---AspGlnAspLys 1360
 Db 5367 GACCGGCTCGGCTGCGGCGCAGCAGCAGATCAATCAATCAATCAATCAATCAATCA 5423
 QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeu-AspThrThrValThrGlnThrAlaGln 1380
 Db 5424 GCTAAGTGTGAAGAGCAGACAGCGCAATCATGTTCTTACAGACAGCGGTGCTGCC 5483
 QY 1380 ValAlaGluLysValIle---GluThrValValIleSerGluThrGlyGln 1396
 Db 5484 TCGTTCGCCAGGTGAGCTGAGCAGCAGCAGCTTCTACTCAGTCTTACGTCCGCTGGGT--- 5541
 QY 1396 userProGluCysValGlyAlaHisIleLeuLeu--- 1406
 Db 5542 -TCTCCGATTTATGCGACTACAGCTGTCTCAGCTTCCCTGCTACCGGCCCACTG 5600
 QY 1407 ---ProAlaGluLysSerSerAlaThrGlyLysIleThrThrLeuGlnHisAlaGluLys 1425
 Db 5601 CAGTTTGTCTCGTCTGCTCCAGAGCCGG--GGTGTCCAGTGGGCGCCCTCCAGAGAGAGAAC 5659

QY 1425 pThrValPro-----LeuGlyProGluSe 1433
Db 5660 GCTTACATGGGACATGTCAGATGAGCTGAGCTGGATGGAACCGCATTTG 5719
QY 1433 rGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeuHisProAs 1453
Db 5720 CAGAGCTGCAGCAGCGCAATCGAGTGTGCCCCACACATCTGAAGACCTGTATCCCTGG 5779
QY 1453 pLeuGlnGlyGluSerAlaSerGlnArgGluArg---SerGluGluGluAspLysPr 1472
Db 5780 AGTCCAGGCTTCCCTGAGCTGGCCACCATCAGATGAGGAGATGAATACTGGAGACC 5839
QY 1472 oAsp-----AlaGlyProAspAlaAspGly 1480
Db 5840 CCAAGAGACCTGCGCCGAGCCAGCATGCGAGCAATCCAGATAGCGGAGGG 5891

RESULT 9
US-08-467-781-3
: Sequence 3, Application US/08467781
: Patent No. 5780596
: GENERAL INFORMATION:
: APPLICANT: TOKUATLY, GARY
: APPLICANT: LIDGARD, GRAHAM P
: TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
: TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TESTA, HURWITZ & THIBEAULT
: STREET: 125 HIGH STREET
: CITY: BOSTON
: STATE: MA
: COUNTRY: USA
: ZIP: 02110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/467,781
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: PITCHER ESQ, EDMUND R
: REGISTRATION NUMBER: 27,829
: REFERENCE/DOCKET NUMBER: MTP-013
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6306 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..6306
: PUBLICATION INFORMATION:
: AUTHORS: COMPTON, DUANE A
: AUTHORS: SZILAK, ILYA
: AUTHORS: CLEVELAND, DON W
: TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
: TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
: TITLE: SEGREGATION OF PROTEINS AT MITOSIS
: JOURNAL: J. Cell Biol.
: VOLUME: 116
: PAGES: 1395-1408
: DATE: 1992
US-08-467-781-3

Alignment Scores:

Pred. No.: 1.98e-19 Length: 6306
Score: 415.00 Matches: 373
Percent Similarity: 35.28% Conservative: 255
Best Local Similarity: 20.96% Mismatches: 663
Query Match: 5.14% Indels: 493
DB: 1 Gaps: 72

US-09-902-432-4 (1-1596) x US-08-467-781-3 (1-6306)

QY 28 GlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAspPro 47
Db 1050 GGCCACTCAGGAGTGGCTAGAGAAGCAGGCCAGCTGGAGAGGAGCTCAGCCGAGCCCT 1109
QY 48 Ala-ThrLys-----LeuProGlnLysAsn-----GlyGlnLeuSerSerVa 61
Db 1110 GCAGGACAAGAAATGCTTTGAAGAGAAGAACGAAATCTTTCAGGGAAAACTTTTCACAGCT 1169
QY 61 1-----AsnGlyValAlaGluGlnGlyAspVal----- 70
Db 1170 GGAAGAACACTTGTCTCCAGCTGCTGAGGATTAACCCACCCAGGAGAGGGGAGGTGTCTGGG 1229
QY 71 -----HisValGlnGluGluAsnGlnGly----- 79
Db 1230 TGATGCTTTCAGCTGGAACCTTGAAGCAAGAGGCGCAGCTCTTCTGCTGCAAAACACAC 1289
QY 80 -----GlnGluGluGluValValAspGluAspValGlyGlnArgGluSerGluAs 96
Db 1290 ACAGCTCCAAAGCAGGCTAGAGATGCTGGAGACTGAGCAGGAGCCAGCAGGAAGCAAGCT 1349
QY 96 pValArgGluLysAspArgValGluGlu----- 105
Db 1350 GCTTGTGTCGCGGGCCACTTTCGAAGAGAAAGCAGCAGCTGTCTAGCTGTATCATCTGA 1409
QY 106 -----MetAlaAlaAsnSerThrAlaValAlaGluAspIleThrLysAspGlyG 121
Db 1410 CCTGCAGAGCTCCATCTCAACCTCAGCCAGGCAAGAGAGAGCTGGAGCGCCTCCCA 1469
QY 121 nGluGluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGluMe 141
Db 1470 GGCTCATGGGCGCGTTCAGCTGCCAGGTGGCTCTCTGACCTCTGAGCTCACCACACT 1529
QY 141 tValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheVa 161
Db 1530 CAATGCCACCATCCAGCAACAGGATCAAGAACTG-----GC 1565
QY 161 lGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLe 181
Db 1566 TGGCCTGAAGCAGCAGGCGCAAGAGAGAGAGCGGCCCTAGCAGACACCTCCAA----- 1620
QY 181 uThrValLysLysAspGluGluGlyAlaGluAlaSerValGlyAlaGlyAspHisG 201
Db 1621 -----CAGCAAGAACAGGCTCCAGGGGCTC---CGCCACCA 1655
QY 201 nGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysG 221
Db 1656 G-----GTGGAGCAGCTAAGCAGTAGCTGAGCAGAGAGAGGAGCAGCAGTTGAAGGA 1706
QY 221 nSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIleProLe 241
Db 1707 GGTAGCGGAGAGCAGGAGGCACT-----AGGCGAGGACCATGCCAGCACTGGCCAC 1760
QY 241 uGlnAlaGluSerAspGlnAlaAla-----GluGluGluAlaLysAspGluGlyGluG 259
Db 1761 TGCTGCAGAGGAGCGAGAGGCTCTTAAAGGAGCGGGATGGCTCTCAAGCAGCTGGA 1820
QY 259 uLysGlnGluLysGluProThrLysSerProGlu-----SerProSerSerProValAs 277
Db 1821 GGCACTGGAGAAGAGAGGCTGCCAAGCTGGAGATTCTGCAGCAGCAACTTCAGGTGTC 1880
QY 277 nSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLy 297
Db 1881 TAATGAAGCCGGGACAGTGGCCAGACCTCATGTGACACAG-----GCCACGCGGA 1931

```
OY 297 slysthrserpheylslyserserlysglasp-----Aspleuqlnthrlagluyl 314
    |||
Db 1932 GAAGCAGAGCTGAGCGGAGAGTGAAGACTCCAGGCTGTGTGAGACCGCGCA 1991
OY 314 sArgylsuglIngluValaspgluglu-----Lysglulysrh 331
    |||
Db 1992 GGAACGACTGAGGCCAGGCCAGGTTCCAGAGCTAGAGTTCCAGCTGGCTGAGCA 2051
OY 331 rgluprolaserglugluIngluProlaglualaspThrAspGlnAlaArgleuserAl 351
    |||
Db 2052 GCAAAAGCACTGAGAAAGAAAGAGGTGCCAGAGAGAACAG-----CTCAGGA 2105
OY 351 aspyrlyglulysValgluleuProlenluaspGlnValGlyAspleuGlualaserse 371
    |||
Db 2106 GCAGCTCCAGGCTCCMAAGAGTCTTGAAGTCCACCAAGGGGCGAGCTTGA----- 2157
OY 371 rgluglulysCyAlaProlenluathrGlulValPheaspGluylMetGluAlaHisgl 391
    |||
Db 2158 -CAGGAGAG-----CGCAGGCTCCAGATGCCCTGGAAGACAGCAGCTTGTCTC 2210
OY 391 nglulValAlaglulVal-----HisValSerThrValglulysrh 405
    |||
Db 2211 TGAGCTGAGGCGAGAGACCGAAGCCTGTGGAGCAGCATTAAGCGGGAACGAAGAGCT 2270
OY 405 rglugluluglInglYlylylylylylylylylylylylylylylylylylylyly 425
    |||
Db 2271 GGAAGAGAGAGAGGCTGGCGCAGAGGGGCTGGAGGCTGATTACTCAGCTTGGGAGGC 2330
OY 425 ygluserleuProlu-----LysleuAlaglulProlglulValProlgl 441
    |||
Db 2331 CCATCAGCTGAGACTGAAGTCTCTCGCGGCGAGCTGGCAGAGGCCATGAGCGTCCAGCA 2390
OY 441 nglulaglul-----ProlaglulgluleuMetLysSerArglulMetCysValserlygl 460
    |||
Db 2391 CACAGCTGAGAGTGAAGTGAAGCAGCTCCCAAGAAAGTGAAGTGGCGGTGACGAGTA 2450
OY 460 yAspHisThrGlnleuThrAspleuserProlglululysThrleuProlHisProlgl 480
    |||
Db 2451 TGAAGTACCCAG-----CAAGAGAGGC-ACAGTATGGCGCATTTCC 2494
OY 480 ugllylValserGluValgluleuMetleuserserGlnluArgllyleValglInglYse 500
    |||
Db 2495 AGAAGACGCTGATGACTTTGA----- 2516
OY 500 rProluylslyleuPheuserSerSerGlyleuLysleuserGlyLyslysglulY 520
    |||
Db 2517 -----GAGGAATGTGAG 2530
OY 520 sglYlysaTarglylylylylylylylylylylylylylylylylylylylylylyly 540
    |||
Db 2531 AGGCCCCGAGGAGCTGACAGAGAGCAAGAGAGAGTGGCAGGCAATGAAATCCACAGCG 2590
OY 540 userProluserAlaAspGlnglulnysglYgluserSerAlaSerSerProlglulP 560
    |||
Db 2591 AGCTCA-----GATTAAGCCGCGCAGCAGAACAACTAGTGAAGCTCCAGCCCAACC 2641
OY 560 ogluglulnthrCysleuGluylsGlyProlenluAlaProlnAspGlyglulala-- 579
    |||
Db 2642 TGGCCAGACACT-----CCAGCAGTCCCAAGAGAAAGAAAGTCCAGGCCCAAACTTG 2695
OY 580 -----GluGlulglY-ThrThrSerAspGlyglulYleYsa 591
    |||
Db 2696 CAGATGACCTTCCTCAGCTGACAGAAAGATGCTCCACACAGCAAGAGAGTGGCCGCT 2755
OY 591 rglulglulylethrrProlProlaserPheylslysmetValThrProlYslysaArgVala 611
    |||
Db 2756 TGGAGCTTGTGCTCCCAAGGCGAGTGAAGAGAAACAGCTCCCGGAGATTAAGTCA 2815
OY 611 rArgProlsergluser-----AsplysGluGlulnlu--LeuGluYsVallyserAla 629
    |||
Db 2816 AGGAGCTCTGAGGCGCAGAGACAGACAGCCGAGTGGCTGGAAGAGCAACAGAGAGCC 2875
OY 629 hrleuserSerThrAspSerThrValserGluMetGlnaspGluValLysThrValglYg 649
```

```
Db 2876 AGTTTCAGACACACAGCAGCGCTGACGCTATGAGCGGAGAGCAGACAGATGGCA 2935
OY 649 lugu-----GlnlysbProlgluln- 655
    |||
Db 2936 ATGAGCTGAAACGGCTGCGGCGCGCTGATGAGAGCCAGGGGCAAGCAGAGAGGAGC 2995
OY 655 ----- 655
Db 2996 GTGGCAGCAGAAAGGAGGTGGCGCGCTGACCCAGAGCGGGCGCTGCCAGGCTG 3055
OY 656 -----P 656
Db 3056 ACCTTGCCCTGAGAAAGCGCCGACAGCAGCTTGAGATGCGCTGCAGAACCCCTCA 3115
OY 656 rOlysaArgValAspThrSerValserThrPglulAlaLeuIleCysValglYserSer 676
    |||
Db 3116 ACGACAGCGTGTGAGATTGCTTACCTGCAAGAGGCACTGGCTCATGCGCTGACGAGAA 3175
OY 676 yLysaArgAlaArglyAlaSerSerSerAspAspGluylYlyProlArgThrleuGlyg 696
    |||
Db 3176 AGGAA-----GGCAAGACACAGAGTT-GGCCAAAGCTTCGTGTCTG 3216
OY 696 lyAspSer-----HisArglaglulglulAlaSerLysAspLysglulala-----GlyT 712
    |||
Db 3217 GAGCAGCCCAAGATTAAGAGCTGAGAGAACTTGCAACCCGTGTAACTGAAGAA 3276
OY 712 hrAspAlaValProla-SerThrGlnglulnAspGlnAlaglInglYserSerSerPro 731
    |||
Db 3277 CA-----GCTGGCTTAAGAAAGAAAGAGCAGCAGCATCTGGCTTCAAGAGCCCA 3324
OY 732 GluProlaglYserProlserGluylglulglYvalSerThrTrpGluSerPheylsArg 751
    |||
Db 3325 TCTGAGGCTGCTGGCAGAGACAGGCAACAGGCCCAAGCTG--GAAGCACTGGCGCA 3381
OY 752 LeuValThrProlArglyLysSerLysSerLysleuGluGluValAspAlaglul----- 768
    |||
Db 3382 GAGGTAGCAAGCTGGAACAGCAATCCAGAAAGCAGCAGAGCAGCTGACAGCTGGA 3441
OY 769 -----AspSerSerValglulnleu 775
    |||
Db 3442 CGCAGCTCGAGGCTGAGCGGCGCTCCGGGCTGAGCGGCAAGTGCCTGAGACTCTG 3501
OY 776 SerThrGluIleGluProlserArggluluserTrpValSerIleLyslePheIlePro 795
    |||
Db 3502 CAGGCGCAGTTGAGAGAGAGGCCAGAGGCTAGGCGCAGTCAAGTGCCTTACCTCG 3561
OY 796 GlYsaArglyLysArgAla-----AspGlysglInglulnAlaThrValgluln 813
    |||
Db 3562 GCCCAACGGAGTGTGCTGCTCCGACCAAGGTAAGACACAGCAGAGGCTGAAGAT 3621
OY 814 SerGlyProlAlglul-----AsnGluAspAspProlserValProlAlaVal 829
    |||
Db 3622 GAGTGAAGGCCAGGTGGCCCGGCGCGGCAAGAGGCTGAGAGAGAAATAGCTCATC 3681
OY 830 ValProlen--SerGluYrAsnAlaValGluArglyLysMetGluAlaglInglYasn 848
    |||
Db 3682 AACACCTTGAGAGAGAGGTGCTCCTCAATCCGAGTCCGAGTCTGAGAGAGAGGG-- 3738
OY 849 ThrGluLeuProlGluLeuGluAlaValIlyrValIserGluGluLeuSerLysThrleu 868
    |||
Db 3739 ---GAGAGCAAGAGATTGAAGCGCTGTGATGGCCGAGTCAAGAGAGAGAGAGCTG 3795
OY 869 ValHisThrValserValAlaValIleaspGlyThrArgAlaValThrSerValgluln 888
    |||
Db 3796 GAGGAGAGCTGGCGCTGCTG-----AGGCAg 3822
OY 889 ArgSerProlserTrpIleSerAlaSerValThrGluProluGluHisThrAlaglYlu 908
    |||
Db 3823 AGACAGCAGCA----- 3834
OY 909 AlaMetProlValglulglulValThrGluYsaPheIleIleAlaglululnthrProlVal 928
```

Db 3835 -----ACAGTGCAGAGCTGCAGAACGCGAGCTCTGCTCGCGGAGG----- 3876
Qy 929 LeuThrGlnThrLeuProGluGlyLysAspAlaHisAspMetValThrSerGlu--- 947
Db 3877 -----AGGTGCAGAGCTCCGGAGGGAGCTGAGAAACAGCGGGTGGCTTCAGAGAAC 3930
Qy 948 -----ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
Db 3931 CTGGCGAGAGCTGACCTCACAGGTGAGCTGCGGAGAGCTGGGCCAAGAAATTGAAG 3990
Qy 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
Db 3991 GCCTGGCAGGAGAGTCTTCAGAAAGACGAGCGCTCTCCACCTCGACGTC----- 4044
Qy 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003
Db 4045 -----GAGCACACACAGCACAGGCC 4065
Qy 1004 ValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaLeuGln 1023
Db 4066 CTGGTGAGTGAGCTGTGCGCAGCTAAGCACTCTGCGCAGAGCTGCGAGCGGAGAGGCC 4125
Qy 1024 AlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
Db 4126 GCTGCCGAGAAACGCCACCGTGGAGCTGGAG-----CAGACAGACGAGGCC 4173
Qy 1044 ThrGlySerLysAlaLeuGlu----- 1050
Db 4174 GCTGGGGAGCTGCGGGCAGAGCTGCTGCGGGCCAGCGGAGCTTGGGGAGCTGATTCTCT 4233
Qy 1051 -----LysValGluGluValGluAspSerGluValLeuAlaSerGluLys--- 1066
Db 4234 CTCCGCGAAGAGTGGCAGAGCAGGAGCGAAGCAAGCTCAGCAGCTGCGGGCAGAGAGGCC 4293
Qy 1066 ----- 1066
Db 4294 AGCTATGCAGACAGCTGACGATGCTGAAGAGCGCATGGCTGTGCGCAGAGGAGAAC 4353
Qy 1066 ----- 1066
Db 4354 CGGGGGCTGGTGAGCGGGCCAACTTGGCGCGGAGTTTCTGGAAAGTGGAGTTGGACCAG 4413
Qy 1067 -----GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeu 1084
Db 4414 GCCCGGGAAGATGTCTCAAGAGTTGGCAGCGCTACGCTGATGCTGAGACCCGCTG 4473
Qy 1085 AlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThr 1104
Db 4474 GCTGAGGTGCGCAGAGCAGACAGAGCACTGCGCGGAGCTGGAGGTG-----ATGACT 4527
Qy 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGln 1124
Db 4528 GCCAAGTATGAGGTGCC-----AAGGTCAAGGTCTCGGAGGAGG 4569
Qy 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
Db 4570 CAGCGGTTCCAGGAAGAGAGCGCAAACTCACTGCCAGGTGGAAGAACTGAGTAAGAAA 4629
Qy 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSer 1162
Db 4630 CTGGTGACTCTGACCAAGCCAGCAGGTGCAGCAGCAGCAAGCTGAAGGTGTCCAGGCT 4689
Qy 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
Db 4690 CAGGGAGCGCAGAGCCAGCAGAGGCCAGCGCTTCAGGCCCGCAGCTGATGACTGCCAA 4749
Qy 1181 AlaAlaThrAlaGlnLysGluGluProSer-----ThrLeu 1192
Db 4750 GCCCAGTTGAGCCAGAGAGGAGCAGCGACTGAGCACTATAAGCTGCAGATGGAGAAAGCC 4809
Qy 1193 ProAsnAsnValProAlaGlnGluGluHisGlyGluGlu-----ProGly 1207
Db 4810 AAAACACATTATGATGCCAAGAGAGCAGAGAACCAAGAGCTGCAGGAGCAGCTGCGGAGC 4869

Qy 1208 ArgAspValLeuLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAla 1227
Db 4870 CTGGAGCAGCTGCAGAAAGGAAACAAAGAGCTCCGAGCTGAAAGCT----- 4914
Qy 1228 LysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLysGlu 1247
Db 4915 ---GAAAGCGCTGGGCATGAGCTACAG-----CAGCTGGGTGAGACCAAGGAG 4962
Qy 1248 GluGlnGluValPheValHis-----SerGlyProAsnSerGlnLysAlaAlaAspVal 1265
Db 4963 GCTGAACAGACCTGCCGCCACCTTACTGCCAGGTGCGAGCTGGAGGCACAGGTTGCC 5022
Qy 1266 ThrTyrrAspSerGluValMetGlyValAlaGlyCysGln-----GluLysGlu 1281
Db 5023 CATGCAGACAGCAGCTTCGAGACCTGGCGCAAAATTCAGGTGGCAACTGATGCTTTAAAG 5082
Qy 1282 SerThrGluValGlnSer-----LeuSer 1289
Db 5083 AGCGGTGAGCCCGAGCTAAGCCCGAGCTGGAGTTCGAGTATTGACAGCTGGATCTGAGC 5142
Qy 1290 LeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGlu 1309
Db 5143 TCGAGAGGGGAGCCCACTCAGTATCACCAGCAGCTGCCTCGT---ACCAGCCAGAC 5199
Qy 1310 GlnValSerGluGluGlyGluGlnGlu-----ThrAlaAlaProGluHisGluGly 1326
Db 5200 GGCACAGCGCTCCGTGGAGAACCCAGCTCACCTATCTCCCA-CGCGCTGCCCCCAAGGT 5258
Qy 1327 ThrTyrrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeu 1346
Db 5259 AGAATCCCTGGAGAGTCTCTACTTCACTCCCATCCCTGCTCGGAGTCAGGCCCCCT--- 5315
Qy 1347 GlySerLeuGlyGlySerProSerLeuPro----- 1356
Db 5316 -----GGAGAGCAGCTGAGCTCCCTGGAGAGCTTCTCTGAGAGCTCGGCTCGTAA 5366
Qy 1357 -----AspGlnAspLys 1360
Db 5367 GACCCGCTCGCTCGTGGCGCAGCAGCATCATCAACATCACCATGACCAAGAA--- 5423
Qy 1361 AlaGlyCysIleGluValGlnValGlnSerLeu-AspThrThrValThrGlnThrAlaGln 1380
Db 5424 GCTAGATGTGGAAGAGCAGCAGCGCAACTCATCTTCTACAGCACGCGTCTGCTCC 5483
Qy 1380 uAlaValGluLysValIle-----GluThrValValIleSerGluThrGlyGln 1396
Db 5484 TGCTTCCAGGTAGCTGCGGAGCCACCTCTCTACTCTAGTCTCTAGCTGCGCTGGGT--- 5541
Qy 1396 uSerProGluCysValGlyAlaHisLeuLeu----- 1406
Db 5542 -TCTCCGATTATGGCAACTCAGCCCTGCTCAGCTTGGCTGCTACCGCCGCCACTCG 5600
Qy 1407 ---ProAlaGluLysSerSerAlaThrGlyHisThrThrLeuGlnHisAlaGluAs 1425
Db 5601 CAGTTCTGCTGCTGCTTCCAGCGCGG-GGTGTCCAGTGGGGCCCTCCAGGAAGGAACA 5659
Qy 1425 pThrValPro-----LeuGlyProGluSe 1433
Db 5660 GCTTCTACATGGGCATTCGCCAGGTAGCCGTGAGCATGCTGGAACCGCATTG 5719
Qy 1433 rGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeuHisProAs 1453
Db 5720 CAGAGCTGCAGCAGCGCAATCAGTGTGCCCCCACTGGAAGACCTGCTATCCCTGG 5779
Qy 1453 pLeuGlnGlyGluIleSerAlaSerGlnArgGluArg-----SerGluGluGluAspLysPr 1472
Db 5780 AGTCCAGGCTTCCCTGAGCTGGGCACCATCATCAGATGAGGAGATGAAAATGGAGACC 5839
Qy 1472 oAsp-----AlaGlyProAspAlaAspGly 1480
Db 5840 CCAAGAGACCTTGGCGCGCAGCAGCATGCAGCCAATCCAGATAGCCGAGGG 5891

```

RESULT 10
US-08-483-924-3
Sequence 3, Application US/08483924
Patent No. 5882876
GENERAL INFORMATION:
APPLICANT: TOUKATLY, GARY
APPLICANT: LUDGARD, GRAHAM P
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08483,924
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6306 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6306
PUBLICATION INFORMATION:
AUTHORS: COMPTON, DUANE A
AUTHORS: SZILAK, ILVA
AUTHORS: CLEVELAND, DON W
TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
TITLE: SEGREGATION OF PROTEINS AT MITOSIS
JOURNAL: J. Cell Biol.
VOLUME: 116
PAGES: 1395-1408
DATE: 1992
US-08-483-924-3

Alignment Scores:
Pred. No.: 1,98e-19 Length: 6306
Score: 415.00 Matches: 373
Percent Similarity: 35.28% Conservative: 255
Best Local Similarity: 20.96% Mismatches: 663
Query Match: 5.14% Indels: 493
Gaps: 72

US-09-902-432-4 (1-1596) x US-08-483-924-3 (1-6306)
QY 28 GYHISGLYProAlaIaGluAlaSerGlyAlaAlaGlyAspProAlaAspPro 47
Db 1050 GGCACCTTCAGAGTGGCTGAGAGAGCAGGCCCTGAGAGAGAGCTCAGCCAGCCCT 1109
QY 48 Ala-ThrIys-----LeuProGlnIysAsn-----GlyGlnIysSerSerVa 61
Db 1110 GCAGGACAAAGAAATGCGCTTGAAGAGAGAGAAATCTTCAGGAGAAACTTTCACAGCT 1169

```

```

QY 61 1-----AsnGlyValAlaGluGlnGlyAspVal----- 70
Db 1170 GGAGAACACTTGTCCAGCTGACGATATACCCAGCCAGAGAGAGGCTGTGGG 1229
QY 71 -----HisValGlnGlnGluAsnGlnGly----- 79
Db 1230 TGATGCTTGCAGCTGGAACCTTGAACCAAGGACCCACTCTGTGCAACACAC 1289
QY 80 -----GlnGlnGlnGluValAlaAspGluAspValGlyGlnArgGluSerGluAs 96
Db 1290 ACAGTCCAGCAGGAGGTAGATGCTGAGACTGAGAGGCGCAGAGAGCAAGCT 1349
QY 96 pValArgGluIysAspArgValGluGlu----- 105
Db 1350 GCTTGCTAGCGGGCCACTTTCGAAAGAGAAAGACAGAGCTGTCTGATCATCTGA 1409
QY 106 -----MetAlaAlaAsnSerThrAlaValAlaGluAspIleThrIysAspGly 121
Db 1410 CCTGCAGAGCTTCATCTCCAACTTCAGCCAGGCCAAGAGAGCTGAGCAGGCTCCCA 1469
QY 121 nGluGluThrSerGluIleIleGlnGlnIleProAlaSerGluAsnValGluGluMe 141
Db 1470 GGCTCATGGGGCCCGTTGACTGCGCCAGTGAGCTCTGCACTCTGAGCTCACCACA 1529
QY 141 tValGlnProAlaGluSerGlnAlaAsnAspValGlyPheIysIysValPheIysPheVa 161
Db 1530 CAATGCCACCATCCAGCAACAGATCAAGAACTG-----GC 1565
QY 161 lGlyPheIysPheThrValIlyAspIysAsnGluIysSerAspThrValGlnLeu 181
Db 1566 TGGCTGAGAGCAGCAGCGCCAAAGAGAGAGCCAGCTAGCAGAGACCTTCCAA----- 1620
QY 181 uThrValIysIysAspGlnGlyGlnGlyAlaGlnAlaSerValGlyAlaGlyAspHisG 201
Db 1621 -----CAGCAAGAAACAGGCTCTCCAGGGCCCTC-----CGCCACCA 1655
QY 201 nGluProSerValGluThrAlaValAlaGlyIysSerAlaSerIysGluSerGluLeu 221
Db 1656 G-----GTGAGAGCAGCTAAGACAGTACCTGAGAGAGAGAGCAGCTTGAAGA 1706
QY 221 nSerThrGluIysGlnGlnGlyThrLeuIysGlnGlnGlnSerThrGlnIlePro 241
Db 1707 GGTAGCGAGAGAGAGAGGAGCACT-----AGCAGAGACCATGCTCCAGCACTGGCAC 1760
QY 241 uGlnAlaGluSerAspGlnAlaAla-----GlnGlnGlnAlaIysAspGlnGlyGlu 259
Db 1761 TGCTGACAGAGAGAGAGAGGCTCTTAAGGAGAGCGGATGGCTCTCAAGCAGCTGGA 1820
QY 259 uIysGlnGlnGluIysGluProThrIysSerProGlu-----SerProSerSerProValAs 277
Db 1821 GGCACCTGAGAGAGAGAGAGGCTCCCAAGCTGAGATTCAGCAGCAACTTCAGTGGC 1880
QY 277 nSerGluThrThrSerSerPheIysIysPhePheThrHisGlyTrpAlaGlyTrpArg 297
Db 1881 TATATAGCCCGGAGCACTGCCACACTCAGTGACACAG-----GCCACAGGGA 1931
QY 297 sIysThrSerPheIysIysSerIysGluAsp-----AspLeuGlnThrAlaGlnGly 314
Db 1932 GAAGCAGAGCTGACCGGAGAGAGTGAAGAACTCCAGGCTGTGTGAGACAGCCGCA 1991
QY 314 sArgIysGlnGlnGlnAlaGlnIysValAspGlnGlnGlu-----LysGlnIysTh 331
Db 1992 GGAAACGATGAGGCGCCAGGCCCAAGTTGACAGAGCTTGCAGCTGCGGCTGAGCA 2051
QY 331 rGluProAlaSerGlnGlnGlnGlnProAlaGluAspThrAspGlnAlaArgLeuSerAl 351
Db 2052 GCAAAAAGCACTGAGAAAGAGAGGTGAGCCAGAGAGAGAGCCAG-----CTCCAGGA 2105
QY 351 aAspTyrgIuIysValGluLeuProLeuGlnIysAspGlnValGlyAspLeuGlnAlaSer 371
Db 2106 GCAGCTCCAGGCGCTCAAAAGAGTCTTGAAGGTTCACCAAGGCGACCTTGAA----- 2157

```

QY 371 rGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisG1 391
Db 2158 -GAGGAGAAG-----CGCAGGGCTGCAGATGCTCCGAAAGCAGCAGCAGGCTGTGATCTC 2210
QY 391 nGluValValAlaGluVal-----HisValSerThrValGluLysTh 405
Db 2211 TGAGCTGAAGCAGAGACCCGAGAGCTGGTGGAGCAGCATAAAGCGGAACGAAGGAGCT 2270
QY 405 rGluGluGluGlnGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrG1 425
Db 2271 GGAAGAAGAGAGAGGCTGGCGCAAGGGCTGGAGGCTCGATTACTGCAGCTTGGGAGGC 2330
QY 425 yGluSerLeuProGlu-----LysLeuAlaGluProGlnGluValProG1 441
Db 2331 CCATCAGCTGAGACTGAAGTCTCGCGCGAGCTGGCAGAGCCATGCTGCCAGCA 2390
QY 441 nGluAlaGlu---ProAlaGluLeuMetLysSerArgGluMetCysValSerGlyG1 460
Db 2391 CACAGCTGAGAGTGAAGTGTGAGCAGCTCGTCAAGAAGTAGTGCCTGGCGGTGACGGTA 2450
QY 460 yAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProG1 480
Db 2451 TGAGGATAGCCAG-----CAAGAGGAGCC-ACAGTATGGCGCCATGTTC 2494
QY 480 uGlyIleValSerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySe 500
Db 2495 AGGAACAGCTGATGACTTTGAA----- 2516
QY 500 rProLeuLysLysLeuPheSerSerSerGlyLeuLysLeuSerGlyLysLysGlnLy 520
Db 2517 -----GGGAATGTGAGA 2530
QY 520 sGlyLysArgGlyGlyGlyGlyAspGluGluProGlyGluTy-rGlnHisIleHisThrG1 540
Db 2531 AGCCCCCAGGAGCTCAGGAGGCAAGGAGAGTGGCGCAGCATAGATCCACAGCG 2590
QY 540 uSerProGluSerAlaAspGluLysGlyGlyLysSerSerAlaSerSerProGluGluPr 560
Db 2591 AGTCCCA-----GATAAGCCGCGCAGCAGCAACAATAGTGTAGCTCCATGCCAAC 2641
QY 560 oGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAla-- 579
Db 2642 TGCCAGAGCACT-----CCAGCAGTCCAGAGAAGGAGTACAGGCCCAAGCTTG 2695
QY 580 -----GluGluGly-ThrThrSerAspGlyGlyLysLysA 591
Db 2696 CAGATGACCTCTCCACTCTGCAGGAAAGATGGTCCACCAGCAAGAGAGTGCCCGCT 2755
QY 591 rGluGlyIleThrProThrProThrAlaSerPheLysLysMetValThrProLysLysArgVala 611
Db 2756 TGGAGACCTTGGTCCGCAAGGCAAGGTGAGCAGCAGGAAACAGAGCTCCCGGGAGTTAGTCA 2815
QY 611 rArgProSerGluSer---AspLysGluGlu---LeuGluLysValLysSerAlaT 629
Db 2816 AGAGCTCGAGGCGAGGAGCAGCAGCCCGAGTGGTGGGAAGAGCAACACAGGAGCGCC 2875
QY 629 hrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyG 649
Db 2876 AGTTCTGCACACACAGCAGCGCTGCAGGCTATGGAGCGGAGGAGCAGACCATGGGCA 2935
QY 649 luGlu-----GlnLysProGluGlu- 655
Db 2936 ATGAGCTGGAACCGCTGCGGGCGCGCTGTATGGAGAGCCAGGGGCGCAGCAGGAGGAGC 2995
QY 655 ----- 655
Db 2996 GTGGGCGAGGAAAGGAGGTGGCGGCTGACCCAGAGCGGGCGCTGCCAGGCTG 3055
QY 656 -----P 656
Db 3056 ACCTTGCCCTGGAGAGGCGCGCAGCAGAGCTTGAGATGCGGCTGCAGAACGCCCTCA 3115
QY 656 roLysArgArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerL 676

Db 3116 ACAGCAGCGTGTGGAGTTGCTTACCTCCAGAGCAGCTGGCTCATGCCCTCAGCGAA 3175
QY 676 yLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyG 696
Db 3176 AGGAA-----GGCAAGGACCAGAGTT-GGCCAAGCTTCTGGTCTG 3216
QY 696 lyAspSer-----HisArgAlaGluGluAlaSerLysAspLysGluAla-----GlyT 712
Db 3217 GAGCAGCCAGATAAAGAGCTGGAGGAACCTTCGCAAAACCTGAACTGAAGGAA 3276
QY 712 hrAspAlaValProAla-SerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerPro 731
Db 3277 CA-----GCTGGCTAAGAAAGAAAGGACGACCATCTCGCTCAGAGCCCAA 3324
QY 732 GluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArg 751
Db 3325 TCTGAGGCTCTGGCAGGACAGAGCCAAAGGCCCCCAAGCTG---GAAGCACTGGGGCA 3381
QY 752 LeuValThrProArgLysLysSerLysSerLysLeuGluLysAlaGlu----- 768
Db 3382 GAGGTGAGCAAGCTGGAACAGCAATGCCAAGACAGCAGCAGCAGCTGACAGCTGGAA 3441
QY 769 -----AspSerSerValGluGlnLeu 775
Db 3442 CGCAGCTCGAGCTGAGCGGCTCCCGGGCTGACGGGACAGTCTCTGGAGACTCTG 3501
QY 776 SerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIlePro 795
Db 3502 CAGGGCCAGTAGAGGAGAGGCCAGGAGCTAGGCCACAGTCAGAGTGCCTTAGCTCG 3561
QY 796 GlyArgLysLysArgAla-----AspGlyLysGlnGlnAlaThrValGluAsp 813
Db 3562 GCCCAACGGAGTGTGCTTCCGCCAACAGGTACAAGACCACACAGAGCTGAAGAT 3621
QY 814 SerGlyProValGluIle-----AsnGluAspAspProAsnValProAlaVal 829
Db 3622 GAGTGAAGGCCAGCTGGCGGCCCGGCCAGAGGCTGAGAGGAAAAATAGCCTCATC 3681
QY 830 ValProLeu---SerGluTyrrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsn 848
Db 3682 AGCAGCTTGAGGAGAGAGTGTCCATCTCGATCCGCGAGTCTGGAGAGGAGGGG--- 3738
QY 849 ThrGluLeuProGlnLeuGlyAlaValTyrrValSerGluGluLeuSerLysThrLeu 868
Db 3739 ---GAGCAAGAGGAGTTGAAGCGGCTGGTGTGATGGCCGAGTCAGAGAAGAGCCAGAAGCTG 3795
QY 869 ValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGlu 888
Db 3796 GAGGAGAGCTGCGCTGCTGC-----AGGCAG 3822
QY 889 ArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGlu 908
Db 3823 AGACCCAGCA----- 3834
QY 909 AlaMetProValGluGluValThrGluLysAspIleAlaGluGluThrProVal 928
Db 3835 -----ACAGTGCAGAGCTGCAGAGCAGCAGCTCTGCTGCGGGAGG----- 3876
QY 929 LeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGlu--- 947
Db 3877 -----AGGTGCAGAGCTCCGGAGGAGGCTGAGAAAACAGCGGTGCTTCAGAGAAC 3930
QY 948 -----ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
Db 3931 CTGCGCAGAGAGTGTACCTCAGAGCTGAGCTGCGGAGAGCTGGCGGAGCTGGCAAGATTGAAG 3990
QY 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
Db 3991 GCGTGCCAGAGAGTGTCTTCAGAAAGAGCAGGCGCTCTCCACCTCGAGCTC----- 4044
QY 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003

Db 4045 -----GAGCACACGACACAGGCC 4065
 Qy 1004 ValGluSerGlyValLeuAspThrGluGluValArgGlnThrGlnAlaIleuGln 1023
 Db 4066 CTGGTAGAGTGAAGCTCTGCTGACCTTAAGCAGCTCTGACGACCTGACGACGACGCC 4125
 Qy 1024 AlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
 Db 4126 GCTGCCGAGAAACGCCCGTAGAGCTGGAG-----CAGAGCAAGAGGCC 4173
 Qy 1044 ThrGlySerLysAlaLeuGlu----- 1050
 Db 4174 GGTGGGAGCATCGGGGACAGCTGCTGGGCGCCAGCGGAGCTTGGGAGCTGATCCT 4233
 Qy 1051 -----LysValGluGluValGluGluAspSerGluValLeuAlaSerGlyLys--- 1066
 Db 4234 CTGGCGAGAAAGTGGCAGAGCAGAGCGAGCAGCTCAGAGCTCGGCGCAGAGAGGCC 4293
 Qy 1066 ----- 1066
 Db 4294 AGCTATGACAGACAGCTGAGCATGCTGAAGAAAGCGCATGCGCTGCTGCGACAGAGAAAC 4353
 Qy 1066 ----- 1066
 Db 4354 CGGGGGCTGGTGAGCGGGCCAACTTGGCCGCGAGCTTCTGGAAGTGGAGCTTGACACAG 4413
 Qy 1067 -----GluLysAspValMetProLysGlyProValGlnGluValaGlyValaGluHisLeu 1084
 Db 4414 GCCCGGAAAGATGATGTCCAGAGTGGCAGCGCTGACGTGATGCTGAGACCCCTGTCG 4473
 Qy 1085 AlaGlnGlySerGluThrGlnAlaThrProGluSerLeuGluValProGluValaThr 1104
 Db 4474 GGTGAGTGCGACCGAGAACACAGACGACTGCCCGGAGCTGAGGTG-----ATGACT 4527
 Qy 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGln 1124
 Db 4528 GCCAAGTATGAGGTCCTC-----AAGTCAAGTCTCTGAGAGAGAG 4569
 Qy 1125 AlaValAlaProGluSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
 Db 4570 CAGCGGTCCAGAGAGAGAGAGAGAACTCAGTGCACAGTGGAGAAACTGATGAGAA 4629
 Qy 1145 LeuAlaAspSerThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSer 1162
 Db 4630 CTGGCTGACTCTGACCAACGACGACGACGACGACGACGACGACGACGACGACGACG 4689
 Qy 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGln 1180
 Db 4690 CAGGAGCGGAGAGCCAGAGAGAGAGCCGCTTCCAGGCCCAAGTGAATGAATGCA 4749
 Qy 1181 AlaAlaThrAlaGlnLysGluGluProSer-----ThrLeu 1192
 Db 4750 GCCCAGTTGAGCCAGAGAGAGAGAGAGAGCTGAGCACTATAGCTGAGATGAGAAAGCC 4809
 Qy 1193 ProAsnAsnValProAlaGlnGluGlnHisGlyGluGlu-----ProGly 1207
 Db 4810 AAAACACATTATGATGCTCAAGAGAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGC 4869
 Qy 1208 ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAla 1227
 Db 4870 CTGGAGCAGCTGACAGAGAGAGAGAGAGAGCTGCGAGCTGAACT----- 4914
 Qy 1228 LysThrGluValGlyGlnGluGluValaAspTrpLeuAspGlyLysValLysGlu 1247
 Db 4915 ---GAACGGCTGGGCGCATGACTACAG-----CAGGCTGGGCTGAAGACCAAGAG 4962
 Qy 1248 GluGlnGluValaPheValHis-----SerGlyProAsnSerGlnLysValaAlaAspVal 1265
 Db 4963 GCTGAACACAGCTGCGCGACCTTACTGACGAGTGGCGGAGCCTGGAGCAGAGTGGCC 5022
 Qy 1266 ThrTyraSpSerGluValMetGlyValaGlyCysGln-----GluLysGlu 1281
 Db 5023 CATGACAGACAGAGCTTGAAGACTGGGCAATTCCAGGTGGCAACTGATGCTTTAAAG 5082

Qy 1282 SerThrGluValGlnSer-----LeuSer 1289
 Db 5083 AGCCCTGAGCCCCAGAGCTTAAGCCCCAGCTGAGCTTGAATTAAGACGCTGATCTGAGC 5142
 Qy 1280 LeuGluGluGlyLysMetGluThrAspValGluLysGlyLysArgGluThrLysProGlu 1309
 Db 5143 TGGAGAGGGGAGCCCACTAGTATACAGAGAAAGTCCCTGCT--ACCAGCCAGAAC 5199
 Qy 1310 GlnValSerGluGluGlyGluGlu-----ThrAlaAlaProGluHisGluGly 1326
 Db 5200 GGACACAGCTCTCCCTGAGAGAACAGAGCTTACCTATCTCCA-GCGCTGCCCCCAAGT 5258
 Qy 1327 ThrTyraGlyLysProValLeuThrLeuAspMetProSerSerGluValaLeu 1346
 Db 5259 AGAATCCCTGAGAGAGCTTCTTCACTCCATCCATCCCTGCTGAGAGTCCGCCCCCT 5315
 Qy 1347 GlySerLeuGlyGlySerProSerLeuPro----- 1356
 Db 5316 -----GAGAGACAGCTGAGCTCCCTGGAGAGAGCTTCTTGACCTGGGCTGTAA 5366
 Qy 1357 -----AspGlnAspLys 1360
 Db 5367 GACCCGCTCCGCTGCGCGGACAGCAGATCATCAACATCAACATCAACAGAA--- 5423
 Qy 1361 AlaGlyCysIleGluValGlnValGlnSerLeu-AspThrValThrGlnThrAlaGlu 1380
 Db 5424 GCTAATGTGAGAGAGCCAGAGCCAGCCACTCATGCTTCAAGCAGCAGCGGTGCTCC 5483
 Qy 1380 AlaValGluLysValIle-----GluThrValValIleSerGluThrGlyGlu 1396
 Db 5484 TGCTTCCAGAGCTGAGCTGCGAGCCAGCTCTCTAAGTCTAAGTCTGAGCTGCGCTGGGT 5541
 Qy 1396 userProGluCysValaGlyAlaHisLeuLeu----- 1406
 Db 5542 -TCTCCGATTAAGCACTGAGCCCTGCTCAGCTTGCCTGAGTACCGCCCAACCACTCG 5600
 Qy 1407 ---ProAlaGluLysSerSerAlaThrGlyGlyHisIleThrLeuGlnHisAlaGluAs 1425
 Db 5601 CAGTTCCTGCTGCTGCTCCAGAGCGCG--GTTGCCAGTGGGCCCCCTCAGAGAAACA 5659
 Qy 1425 pThrValPro-----LeuGlyProGluLeu 1433
 Db 5660 GCTTTCATGAGGCACTTGCACAGATGAGCTGAGCAGCTGAGTACTGAAACCCATTG 5719
 Qy 1433 GlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeuHisAspAs 1453
 Db 5720 CAGAGCTGACAGAGCGCAATGAGTGTCCCCCACTGAAGACCTGATCCCTGG 5779
 Qy 1453 pLeuGlnGlyLysLeuSerAlaSerGlnArgGluArg--SerGluGluGluAspLysPr 1472
 Db 5780 AGTCCAGGCTTCCCTGAGCCTGGCAGCAACATCAAGAGAGAGTGAAGAACTGAGAAC 5839
 Qy 1472 GASP-----AlaGlyProAspAlaAspGly 1480
 Db 5840 CCAAGAGACCTGCGCGGACAGCAGATGACCAATCCAGATAGCCGAGGG 5891

RESULT 11
 US-08-195-487-3
 Sequence 3, Application US/08195487
 Patent No. 5783403
 GENERAL INFORMATION:
 APPLICANT: TONKATLY, GARY
 APPLICANT: LIDGARD, GRAHAM P
 TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
 TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESS: TESTA HURWITZ & THIBEAULT
 STREET: 53 STATE STREET
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,487
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/901,701
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6306 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6306
PUBLICATION INFORMATION:
AUTHORS: COMPTON, DUANE A
AUTHORS: SZILAK, ILLYA
AUTHORS: CLEVELAND, DON W
TITLE: PRIMARY STRUCTURE OF OF NUMA, AN INTRANUCLEAR
TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
TITLE: SEGREGATION OF PROTEINS AT MITOSIS
JOURNAL: JOURNAL OF CELL BIOLOGY
VOLUME: 116
PAGES: 1395-1408
DATE: MAR-1992
US-08-195-487-3

Alignment Scores:
Pred. No.: 3,2e-19 Length: 6306
Score: 412.00 Matches: 372
Percent Similarity: 35.28% Conservative: 256
Best Local Similarity: 20.90% Mismatches: 663
Query Match: 5.10% Indels: 493
DB: 1 Gaps: 72

US-09-902-432-4 (1-1596) x US-08-195-487-3 (1-6306)

QY 28 GlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAspAlaAspPro 47
Db 1050 GGCACCTCAGAGTGGCTAGAGAGAGAGGCGCCAGCTGGAGAGGAGCTCAGCGCAGCCT 1109
QY 48 Ala-ThrLys-----LeuProGlnLysAsn-----GlyGlnLeuSerSerVa 61
Db 1110 GCAGGACAGAAATGCTTGNAGAGAGAGACGAATCTCTTCAGGGAATCTTCACAGCT 1169
QY 61 l-----AsnGlyValAlaGluGlnGlyAspVal----- 70
Db 1170 GGAAGAACCTTGCCAGCTGCAGGATACCCACCCAGGAGAGGCGAGGTGCTGGG 1229
QY 71 -----HisValGlnGluGluAsnGlnGly----- 79
Db 1230 TGATGTCTTGAGCTGGAAACCTTTGAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1289
QY 80 -----GlnGluGluGluValValAspGluAspValGlyGlnArgGluSerGluAs 96
Db 1290 ACAGCTCAAGCCAGGCTAGAGATGCTGGAGACTGAGCGAGGCCAGCAGGAGCAAGCT 1349

QY 96 pValArgGluLysAspArgValGluGlu----- 105
Db 1350 GCTTGTGAGCGGGGCCACTTCGAAGAAGAAAGACAGCAGCTCTGTAGCTGTGATCACTGA 1409
QY 106 -----MetAlaAlaAsnSerThrAlaValAlaGluAspIleThrLysAspGlyG1 121
Db 1410 CTTGCAGAGCTCCATCTCCAACTCAGCCAGGCGCAAGAGAGAGCTGGAGCAGCCTCCCA 1469
QY 121 nGluGluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGluMe 141
Db 1470 GGCTCATGGGGCCGGTTGACTGCCAGGTGGCTCTCTGACCTCTGAGCTCACCACACT 1529
QY 141 tValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheVa 161
Db 1530 CAATGCCACCATCCAGCAACAGGATCAAGAACTG-----GC 1565
QY 161 lGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLe 181
Db 1566 TGGCCTGACAGCAGCGCCAAAGAGAGAGAGCGCCAGCTAGCACAGACCTCCAA----- 1620
QY 181 uThrValLysLysAspGluGlyAlaGluAlaSerValGlyAlaGlyAspHisG1 201
Db 1621 -----CAGCAAGAACAAGCGCTCCAGGGCCTC-----CGCCACCA 1655
QY 201 nGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysG1 221
Db 1656 G-----GTGGAGCAGCTAAGCAGTAGCTGGAAGCAGAGAGGAGCAGCAGTTGAAGGA 1706
QY 221 nSerThrGluLysGlnGlyThrLeuLysGlnGlnSerSerThrGluIleProLe 241
Db 1707 GTAGCGGAGAGCAGGAGCAACT-----AGCAGGACCATTGCCAGCAACTTGGCCAC 1760
QY 241 uGlnAlaGluSerAspGlnAlaAla-----GluGluGluAlaLysAspGluGlyGluG1 259
Db 1761 TGCTGCAGAGGAGCGAGAGGCTCTTAAAGGAGCGGATGCGCTCTCAAGCAGCTGGA 1820
QY 259 uLysGlnGluLysGluProThrLysSerProGlu-----SerProSerSerProValas 277
Db 1821 GGCACCTGGAGAGAGAGAGGCTGCCAAGCTGGAGATTCTGCAGCAGCAACTTTCAGGTGGC 1880
QY 277 nSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLy 297
Db 1881 TAATGAAGCCCGGACAGCTGCCAGACCTCAGTGCACACAG-----GCCAGCGCGGA 1931
QY 297 sLysThrSerPheLysLysSerLysGluAsp-----AspLeuGluThrAlaGluLy 314
Db 1932 GAAGGCAGAGCTCAGCGGAGAGTGGAGGAACCTCCAGGCTGTGTGAGACAGCCGCCCA 1991
QY 314 sArgLysGluGlnGluAlaGluLysValAspGluGluGlu-----LysGluLysTh 331
Db 1992 GGAACAGCATGAGGCCCGCCAGGCTTGCAGAGCTAGAGTTGCAGCTGCGGTCTGAGCA 2051
QY 331 rGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAl 351
Db 2052 GCAAAAGCAACTTGAGAAAGAGGTTGGCCCGCAGGAGAGAGGAGCAG-----CTCCAGGA 2105
QY 351 aAspTyrGluLysValGluLeuProLysGlnValGlyAspLeuGluAlaSerSe 371
Db 2106 GCAGCTCCAGGCCCTCAAGAGTCTTGAAGGTCCACCAAGGCGAGCTTGAA----- 2157
QY 371 rGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisG1 391
Db 2158 -GAGGAGNAG-----CGCAGGGCTGCAGATGCTCCCTGGAAGAGCAGCAGCGGTGTATCTC 2210
QY 391 nGluValValAlaGluVal-----HisValSerThrValGluLysTh 405
Db 2211 TGAGCTGAAGGSCAGAGACCCGAAAGCTGTGGAGCAGCATTAAGCGGGAACAAAGAGGCT 2270
QY 405 rGluGluGluGlnGlyGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrG1 425
Db 2271 GGAAGAGAGAGAGGCTGGCGCGAAGGGGCTGGAGGCTCGATTACTGACGCTTGGGAGGC 2330
QY 425 yGluSerLeuProProGlu-----LysLeuAlaGluProGlnGluValProG1 441

```

Db 2331 CCATCAGCTGAGACTGAATCTCTGCGGGAGCTGGCAGAGGCCATGCTGCCACAGA 2390
Qy 441 ngluaaglu---ProLaIgluIleuMetLysSerArgIleuMetCysValSerGly 460
Db 2391 CACAGCTGAGAGTGATGTGGAGCAGCTCGTCAAGAAAGTAGTGTGCTGGCGTGAAGGGGA 2450
Qy 460 yAspHisThrGluLeuThrAspLeuSerProGluIuLysThrLeuProLysHisProG 480
Db 2451 TGAGGATGAGCCAG-----CAAGAGGAGGC-ACAGTATGGCGCGCATGTTC 2494
Qy 480 uGlyIleValSerGluValGluMetLeuSerSerGlnIuArgIleLysValGlnGlySe 500
Db 2495 AGGAACACAGCTGATGACTTTGAA----- 2516
Qy 500 rProleuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysGlnLys 520
Db 2517 -----GGAGGAATGTGAGA 2530
Qy 520 sGlyLysArgGlyGlyGlyLysPheGluIuProGlyIuTrGlnHisIleHisThrG 540
Db 2531 AGGCCCCGAGAGCTGGCAGAGGCAAGAGAGAGTGGCAGCATAGATCCACAGCG 2590
Qy 540 userProGluSerAlaAspGluGlnLysGlyLysSerAlaSerSerProGluIuPr 560
Db 2591 AGCTCCA-----GATTAAGCCGCGCAGCAGCAACAACTGAGCTCAATGCCAAC 2641
Qy 560 oGluIuThrThrCysLeuGluLysGlyProleuGluIaProGlnAspGlyGluIa-- 579
Db 2642 TGGCCAGAGCACT-----CCAGCAGGTCTCAAGAGAGAACTCAGGCCCGCAAGCTTG 2695
Qy 580 -----GluGluGly-ThrThrSerAspGlyLysLysA 591
Db 2696 CAGATGACCTCTCCACTCTGCAGAAAGATGCTGCCACAGCAAGAGAGTGGCCGCT 2755
Qy 591 rGluGluGlyIleThrProTrpAlaSerPheLysLysMetAlaThrProLysLysArgVala 611
Db 2756 TGGAAACCTTGCTGCCGCAAGCAGGTGAGCAGAGGAAACAGCTCCCGGAGTAGTGA 2815
Qy 611 rGArgProSerGluSer---AspLysGluGluGlu---LeuGluLysValLysSerAlaT 629
Db 2816 AGAGACCTTCGCGAGGAGGAGCAGACAGCCGAGTGCCTGGAAGCAACAGGAGCGCC 2875
Qy 629 hrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyG 649
Db 2876 AGTTCAGAGCAGACAGGCGCTGCAGGCTATGAGCGGAGGAGGAGCAGATGGGCA 2935
Qy 649 luGlu-----GlnLysProGluGlu- 655
Db 2936 ATGAGCTGGAAAGCTGCGCGCGCGCTGATGAGAGAGCCAGGGCAGCAGAGAGAGC 2995
Qy 655 ----- 655
Db 2996 GTGGGACAGCAAGAGGAGGTGGCGGCTGACCAGAGAGCGGCGCTGCCAGGCTG 3055
Qy 656 -----P 656
Db 3056 ACCTTGCCCTGGAGAAAGCGCCAGAGCAGAGCTTGAGATGCGGCTGCAGAAAGCCCTCA 3115
Qy 656 rGlyArgArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerL 676
Db 3116 ACGACAGAGGTGTGAGTTCCTACCTGCAAGAGGACATGGCTCATGCCCTGACGGAAA 3175
Qy 676 yLysArgArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyG 696
Db 3176 AGGAA-----GCCAAGACCAAGAGTT-GGCCAAGCTTGTGTTCTG 3216
Qy 696 lYAspSer-----HisArgAlaGluGluAlaSerLysAspLysGluAla-----GlyT 712
Db 3217 GAGGAGGCCCAATAAAGAGCTGGAGAGAACTTGGCAAAACCGTGAAGCAACTGAAGGAA 3276
Qy 712 hrAspAlaValProAla-SerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerPro 731

```

```

Db 3277 CA-----GCTGGCTAAGAAAGAAAGAGCAGCATGCTGGCTCAGAGCCCA 3324
Qy 732 GluProAlaGlySerProSerGluGlyGluIleValSerThrTrpGluSerPheLysArg 751
Db 3325 TCTGAGGCTGCTGGCAGAGGACAGAGCCAAAGAGCCCAAGCTG---GAAGCAGCTGGGCA 3381
Qy 752 LeuValThrProAlaGlyLysLysSerLysSerLysLeuGluGluLysValIleGlu----- 768
Db 3382 GAGGTGAGCAAGCTGGAAACAGCAATGCCAAGAGAGCAGGAGGCTGAGACGCTGGAA 3441
Qy 769 -----AspSerSerValGluGluLeu 775
Db 3442 CGCAGCTCGAGGCTGAGCGGGCTCCCGGCTGAGCGGAGCACTGCTCTGGAAGACTTG 3501
Qy 776 SerThrGluIleGluProSerArgGluLeuSerTrpValSerIleLysLysPheIlePro 795
Db 3502 CAGGCGCAGTTAGAGAGAAAGCCAGAGAGCTAGAGGACAGTCAAGAGTGTGAGCTCG 3561
Qy 796 GlyArgArgLysLysArgAla-----AspGlyLysGlnGluGlnAlaThrValGluAsp 813
Db 3562 GCCAAGCGGAGTTGCTGCTGCTCCGACCAAGGTACAAAGCAGCAGAGCTGAAGAT 3621
Qy 814 SerGlyProValGluIle-----AsnGluAspAspProAsnValProAlaVal 829
Db 3622 GAGTGAAGGCCCAGTGGCCCGGGCGCGGCAAGAGGCTGAGAGAAATAAGCTCATC 3681
Qy 830 ValProLeu---SerGlyTrpAsnAlaValGluArgGlyLysMetGluAlaGlnGlyAsn 848
Db 3682 AGCAGCTTGAGAGAGAGGTGCTCATCTGAATCGCCAGGTCCTGGAGAAAGAGGG--- 3738
Qy 849 ThrGluLeuProGlnLeuLeuGlyAlaValIleValSerGluGluLeuSerLysThrLeu 868
Db 3739 ---GAGACCAAGAGTTGAACCGCTGTGTATGGCCGAGTCAAGAGAGAGCAGAGCTG 3795
Qy 869 ValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGlu 888
Db 3796 GAGAGAGCTGCGCTGCTGTC-----AGGCAG 3822
Qy 889 ArgSerProSerTrpIleSerAlaSerValThrGluProleuGluHisThrAlaGlyL 908
Db 3823 AGACAGCAGCA----- 3834
Qy 909 AlaMetProProValGluGluValThrGluLysAspIleIleAlaGluIuThrProVal 928
Db 3835 -----ACAGTGCAGAGCTCGAGAACCGACCTGCTCGCGGAG- 3876
Qy 929 LeuThrGlnThrLeuProGluGlyLysAspAlaHisAspMetValThrSerLys--- 947
Db 3877 -----AGGTGCAAGCTCCGGAGGAGAGGCTGAGAAACAGCGGTGGCTTCAGAGAAC 3930
Qy 948 -----ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
Db 3931 CTGGCGCAGAGAGCTGACCTCAAGGCTGAGCGTGGAGAGCTGGCGCCAGAAATTGAAG 3990
Qy 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
Db 3991 GCGTGGCAGAGAAAGTTCTTCAGAAAGAGAGCGCTTCACCTCCAGCTC----- 4044
Qy 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003
Db 4045 -----GAGCAGACCAAGCAGAGGCC 4065
Qy 1004 ValGluSerGlyValLeuAspThrGluGluGluGluGluGluGluGluGluGluGluG 1023
Db 4066 CTGCTGATGAGCTGCTCCAGAGCTAAGACCTTGCAGAGAGCTGACAGCGGCGAGAGGCC 4125
Qy 1024 AlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
Db 4126 GCTGCCAGAAACGCCACCGAGGAGAGCTGGAG-----CAGAGCAAGCAGAGCC 4173
Qy 1044 ThrGlySerLysAlaLeuGlu----- 1050
Db 4174 GCTGGGGAGCTGGCGGAGAGCTGCTGCGGGCCAGCGGAGCTTGGGAGCTGATTCT 4233

```

```
QY 1051 -----LysValGluGluValGluAspSerGluValLeuAlaSerGluLys--- 1066
Db 4234 CTGGGAGAAAGTGGCAGAGCAGGACCGAACAGCTCAGCAGCTGCGGAGAGAGGCC 4293
QY 1066 ----- 1066
Db 4294 AGCTATCAGACAGCTGACATGCTGAAGAGCGCATGGCTGTGGCAGAGGAGAAC 4353
QY 1066 ----- 1066
Db 4354 CGGGGGCTGGGTAGCGGGCCAACTTTGGCGCGCAGTTTCTGGAAGTGGAGTGGACCA 4413
QY 1067 -----GluLysAspValMetProLysGlyProValGlnGluAlaGlyHisLeu 1084
Db 4414 GCCCGGAAAAGTATGTCGAAGAGTTGGCAGCGCTACGTGCTGATGCTGAGACCCGCTG 4473
QY 1085 AlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThr 1104
Db 4474 GCTGAGGTGACGAGNAGCAGACAGACACTGCCGCGGAGCTGGAGTG-----ATGACT 4527
QY 1105 AlaAspValAspHisValAlaThrCysGlnValLeuLysLeuGlnLeuMetGluGln 1124
Db 4528 GCCAAGTATCAGGTGCC-----AAGTCAAGGTCTCTGGAGGAGG 4569
QY 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
Db 4570 CACCGGTTCCAGGAAGAGAGCGCAAACTCACTGCCAGGTGGAAGAACTGAGTAAGAAA 4629
QY 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSer 1162
Db 4630 CTGGCTGACTCTCACCAGCCAGCAGAGTGCAGCAGAGCAAGCTGAAGGCTGTCCAGGCT 4689
QY 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu 1180
Db 4690 CAGGGAGGCGAGAGCCAGCAGGAGGCCAGCGCTTCAGGCCCCAGCTGAATGAACTCAA 4749
QY 1181 AlaAlaThrAlaGlnLysGluGluProSer-----ThrLeu 1192
Db 4750 GCCCAGTTGAGCCAGAGAGGAGCAGCGACTGATTAAGCTGCAGATGGAGAAAGCC 4809
QY 1193 ProAsnValProAlaGlnGluGluHisGlyGluGlu-----ProGly 1207
Db 4810 AAACACATTATGATGCCAAGAGAGCAGCAGAACACAGAGCTGCAGGAGCAGTCCGGAGC 4869
QY 1208 ArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAla 1227
Db 4870 CTGGAGCAGCTGCAGAGGAAACAAAGAGCTGCGAGCTGAAGCT----- 4914
QY 1228 LysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysValLysGlu 1247
Db 4915 --GAACGGCTGGGCCATGAGCTACAG-----CAGGCTGGCTGAAAGACCAAGGAG 4962
QY 1248 GluGlnGluValPheValHis-----SerGlyProAsnSerGlnLysAlaAlaAspVal 1265
Db 4963 GCTGACACAGACTCCGCCCACTTACTGCTCCAGGTGGCAGCTGGAGGCACAGGTGCC 5022
QY 1266 ThrTyrAspSerGluValMetGlyValAlaGlyCysGln-----GluLysGlu 1281
Db 5023 CATGCAGACCAGCAGCTTCAGACCTGGGCAAAATTCAGGTGGCACTGATGCTTTAAAG 5082
QY 1282 SerThrGluValGlnSer-----LeuSer 1289
Db 5083 AGCCGTGAGCCCGGCTAAGCCCGCAGCTGGATTGACAGCTGGATCTGAGC 5142
QY 1290 LeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgGluThrLysProGlu 1309
Db 5143 TGGAGAGGAGGAGCCCACTCACTATCACCAGCAAGCTGCTCGT---ACCCAGCCAGAC 5199
QY 1310 GlnValSerGluGlyGluGlnGlu-----ThrAlaAlaProGluHisGluGly 1326
Db 5200 GGCACACAGCTCCTGGAGAACCAAGCCTACCTATCTCCCA-GCGCTGCCCCCAGGTT 5258
```

```
QY 1327 ThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArgGlyLysAlaLeu 1346
Db 5259 AGAATCCCTGGAGAGTCTCTACTTCACTCCCATCCCTGCTCGGAGTCAAGCCCCCT--- 5315
QY 1347 GlySerLeuGlyLysSerProSerLeuPro----- 1356
Db 5316 -----GGAGAGCAGCTGGACTCCCTGGGAGAGCGTCTTCTGAGCTCGGGTCTGTAA 5366
QY 1357 -----AspGlnAspLys 1360
Db 5367 GACCCGCTCGCTCGTGGCGCACCGAGATCATCAATCACCATCACCACCAAGAA--- 5423
QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeu-AspThrThrValThrGlnThrAlaGln 1380
Db 5424 GCTAGATGCGAAGAGCAGCAGCGCCCACTCATCTGTTCTACAGCAGCGGGTCTGCTCC 5483
QY 1380 uAlaValGluLysValIle-----GluThrValValIleSerGluThrGlyGln 1396
Db 5484 TGTTTCCAGGCTAGCTGGAGCCACCTCTCTACTAGTCTCTAGCTCGCTGGGT-- 5541
QY 1396 userProGluCysValGlyAlaHisLeuLeu----- 1406
Db 5542 -TCTCCGATTATGCAACTCAGCCCTGCTCAGCTTGGCTACCGCCCACTCG 5600
QY 1407 -----ProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAs 1425
Db 5601 CAGTTCTGCTGCTGCTTCCAGCGCGG-GGTGTCCAGTGGGGCCCTCCAGGAAGGAACA 5659
QY 1425 pThrValPro-----LeuGlyProGluSe 1433
Db 5660 GCTTCTACATGGCCTTGCAGGATGACCTGAGCAGCTGGATGACTGGAACCGCATTC 5719
QY 1433 rGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeuHisProAs 1453
Db 5720 CAGAGCTGCAGCAGCGCAATCGAGTGTGCCCCCACTGAAGACCTGCTATCCCTCG 5779
QY 1453 pLeuGlnGlyGluIleSerAlaSerGlnArgGluArg---SerGluGluGluAspLysPr 1472
Db 5780 AGTCCAGGCTTCCCTGAGCCTGGGCACCATCACAGATGAGGAGATGAATACTGGAGACC 5839
QY 1472 oAsp-----AlaGlyProAspAlaAspGly 1480
Db 5840 CCAAGAGACCTTGCAGCGGAGCAGCATGCGCAATTCAGATAGCCGAGGG 5891

RESULT 12
PCT-US93-06160-3
; Sequence 3, Application PC/TUS9306160
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06160
; FILING DATE: 19930621
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ. EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
```

	TELEPHONE: 617/248-7000	
	TELEFAX: 617/248-7100	
	INFORMATION FOR SEQ ID NO: 3:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 6306 base pairs	
	TYPE: nucleic acid	
	STRANDEDNESS: single	
	TOPOLOGY: linear	
	MOLECULE TYPE: cDNA	
	FEATURE:	
	NAME/KEY: CDS	
	LOCATION: 1..6306	
	PUBLICATION INFORMATION:	
	AUTHORS: COMPTON, DUANE A	
	AUTHORS: SZILAK, ILIYA	
	AUTHORS: CLEVELAND, DON W	
	TITLE: PRIMARY STRUCTURE OF OF NUMA, AN INTRANUCLEAR	
	TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR	
	TITLE: SEGREGATION OF PROTEINS AT MITOSIS	
	JOURNAL: JOURNAL OF CELL BIOLOGY	
	VOLUME: 116	
	PAGES: 1395-1408	
	DATE: MAR-1992	
	PCT-US93-06160-3	
	Alignment Scores:	
	Pred. No.: 3.2e-19	Length: 6306
	Score: 412.00	Matches: 372
	Percent Similarity: 35.28%	Conservative: 256
	Best local Similarity: 20.90%	Mismatches: 663
	Query Match: 5.10%	Indels: 493
	DB: 5	Gaps: 72
US-09-902-432-4 (1-1596) x PCT-US93-06160-3 (1-6306)		
Qy 28	GLYHISGLYPRQALAAIGLUALASERGIYALALAGLYASP	ProAlaAspProAlaAsp

	DB 1050	GSCCACTCAGAGGTGCTTAAGAAAGAGGCCACGTGAGAAAGAGCTCAGCGCAGCCCT
		11099
Qy 48	ALA-ThrIys-----LeuProGlnIysAsn-----GlyGlnLeuSerSerVa 61	

	DB 1110	GCAGACACAAAGAAATGCTTTAAAGAGAAACGAATCTTCAGGGGAAACTTCACACT
		11699
Qy 61	I-----AsnGlyValAlaGlnGlnGlyAspVal----- 70	

	DB 1170	GAAAGAACACTTGTCCACAGCTGCAGATAACCCACCCACAGAGAAAGGCGAGGTCTGGG
		12299
Qy 71	-----HisValGlnGlnGlnGlnAsnGlnGlnGly----- 79	

	DB 1230	TGATGCTTTCAGAGCTGGAAACCTTGAAGCAAGAGGAGCAGCCACTCTTGCTGCAACACAC
		12899
Qy 80	-----GlnGlnGlnGlnValValaAspGluAspValGlyGlnArgGlnSerGluAs 96	

	DB 1290	ACAGCTCCAGCCAGCGGTAGAGATGCTGGAGACTGACAGCAGGCGCAGACAGAACCAACT
		13499
Qy 96	PValArgGlnIlysAspArgValGlnGlu----- 105	

	DB 1350	GCTTGCTAGCGGGGCCACTTCCGAAGAAAGAAAGCAGACGTGTCTAGCTGATCACTGA
		14099
Qy 106	-----MetAlaIAsnSerThrAlaValaGluAspLleThrIlysAspGlyG 121	

	DB 1410	CCTGCAGAGCTCCATCTCCCAACTCAGACCGCCAGCAAGAAAGCTGAGACAGGCTCCCA
		14699
Qy 121	nGlnGlnThrSerGlnIleIleGlnGlnIleProAlaSerGlnAsnAsnValGlnGlu 141	

	DB 1470	GAGCTATGAGGGGCCGGTGTGACTGCCACAGGTGGCTCTTGACCTCTGAGCTACACACACT
		15299
Qy 141	tValGlnProAlaGlnSerGlnAlaAsnAspValGlyPheIysValPheIysPheVa 161	

	DB 1530	CAATGCCACCATCCAGCAACAGGATCAAGAACTG-----GC
		15699
Qy 161	IgIyPheIysPheThrValIysIysAspIysAsnGlnIysSerAspThrValGlnLeu 181	

Db	1566	TGCGCTGAACGACAGAGCCAAAGAAAGAGAGCCGACGCTAGACAGACCTTCCAA-----	1620
Qy	181	uThrValLysLysAspGlnGlyGluGlyValAlaGluAlaSerValGlyAlaGlyAspHisGly	201
Db	1621	-----CAGCAAGAAAGAGCCCTTCCAGAGGCCCTCC-----CGCCACCA	1655
Qy	201	nGluProSerValGluThrAlaValGlyLysSerAlaSerLysGlySerGlyLeuLysGly	221
Db	1656	G-----GTGAGACAGCTAAGACATGAGCTCTGAAGCAAGAAAGGACGACGCTTGAAGGA	1706
Qy	221	nSerThrGluLysGlnGluGlyThrLeuLysGlnGlnInserSerThrGluLeuProLe	241
Db	1707	GGTAGCGGAGAAACAGAGAGGACAAT-----AGCAGAGACATGCTCCAGCAACTGGCCAC	1760
Qy	241	uGlnAlaGlySerAspGlnAlaAla-----GluGlnGluAlaLysAspGlnGlyGluGly	259
Db	1761	TGCTGCAGAGAGACGAGAGCCCTCTTAAAGGAGCGGAGTGGCTCTCAACAGCTGGA	1820
Qy	259	uLysGlnGluLysGluProThrLysSerProGlu-----SerProSerSerProValAs	277
Db	1821	GGCATGGAAGAAAGAAAGGCTGCCAGCTGGAGATTTCTGCAGACGCAATTCAGGTGC	1880
Qy	277	nSerGluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLys	297
Db	1881	TAAAGAAAGCCCGGAGACAGTCCACAGACCTCAGTGAACAG-----GGCCAGCGGGA	1931
Qy	297	sLysThrSerPheLysLysSerLysGluAsp-----AspLeuGluThrAlaGluLys	314
Db	1932	GAAAGCAGACCTGAGCCGCGAAGGTGAGAGATCCAGGCTGTGTGTGAGACAGCCGCA	1991
Qy	314	sArgLysGlnGlnGluAlaGluLysValAspGlnGluGlu-----LysGlnLysLys	331
Db	1992	GAAACAGCATGAGCCCGACGCCCAGGCTTCCAGAGCTAGAGTTGACAGCTGCGGTGAGCA	2051
Qy	331	rGluProAlaSerGlnGluGlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAl	351
Db	2052	GCAAAAAGCACTGAGAAAGAAAGAGGTGGCCGACAGAAAGAGCCAG-----CTCCAGA	2105
Qy	351	aAspTrpGluLysValGluLeuProLeuGlnLysAspGlnValGlyAspLeuGluAlaSer	371
Db	2106	GCAGCTCCAGGCGCTCAAAAGATGCTTGAAGGTCAACAAGGCGACAGCTTTGAA-----	2157
Qy	371	rGlnGluLysCysAlaLeuProLeuAlaThrGluValPheAspGlyLysMetGluAlaHisGly	391
Db	2158	-GAGAAAG-----CGAGGGCTCAAGATGCTCTGAAAGCGACGAGCGTGTATCTC	2210
Qy	391	nGlyValValAlaGluVal-----HisValSerThrValGluLysLys	405
Db	2211	TGAGCTGAAGGCAAGACCCGAAAGCTGTGTGAGACGCTAAGCGGAAACGAAAGAGACT	2270
Qy	405	rGlnGluGlnGlnGlyGlyGlyGlyGluAlaGluGlnGlyValValGluGlyThrGly	425
Db	2271	GGAAGAAAGAGGGCTGGGCGCAAGGGGCTGGAGGCTCGATTACTGACAGCTTGGGGAGGC	2330
Qy	425	yGluSerLeuProProGlu-----LysLeuAlaGluProGlnGluValProGly	441
Db	2331	CCATCAGGCTGAGACTGAATCTCTGGCGGGAGAGCTGGAGAGGCGCATAGCTGCCACGA	2390
Qy	441	nGluAlaGluProAlaGlnGluLeuMetLysSerArgLysMetCysValSerGlyGly	460
Db	2391	CACAGCTGAGAGTGAGTGATGACAGCTCTGTCAAAGAAAGTACGTGCTGGCGTGAAGGGTA	2450
Qy	460	yAspHisThrGlnLeuThrAspLeuSerProGlnGluLysThrLeuProLysHisProGly	480
Db	2451	TGAGGATAGCCAG-----CAAGAGAGGC-ACAGTATGGCGCCATGTCC	2494
Qy	480	uGlyTyrLeuValSerGluValGluMetLeuSerGlnGluArgGlyLeuValGlnGlySer	500
Db	2495	AGGAACAGCTGATCATCTTTAA-----	2516
Qy	500	rProLeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlnLys	520
Db	2517	-----GAGGAAATGTGAGA	2530

QY 520 sGlyLysArgGlyGlyGlyAspGluGluProGlyGluTyrGlnHisIleHisThrG1 540
Db 2531 AGGCCCCAGGAGCTGCAGGAGCAAGAGAGAGGTGGCAGGATAGAAATCCACAGCG 2590
QY 540 uSerProGluSerAlaAspGluGlnLysGlyGlyGlySerAlaSerSerProGluGluPr 560
Db 2591 AGCTCCA-----GATAGCCGGCAGCAGAAACAATAGCTGAGCTCATGCCAAC 2641
QY 560 oGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAla-- 579
Db 2642 TGGCCAGAGCACT-----CCAGCAGGTCCCAAGAGAGGAAGTCAAGGCCAGAGCTTG 2695
QY 580 -----GluGluGly-ThrThrSerAspGlyGluLysLysA 591
Db 2696 CAGATGACCTCTCCACTCTGCAGGAAAGATGGCTGCCACCAAGAGAGTGGCCGCT 2755
QY 591 xGluGluIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValA 611
Db 2756 TGGAGACTTGTCCGCAAGCAGGTGAGCAGCAGGAGAAACAGCCTCCCGGAGTTAGTCA 2815
QY 611 rgArgProSerGluSer---AspLysGluGluGlu---LeuGluLysValLysSerAlaT 629
Db 2816 AGGACCTCGAGGCGAGGAGACAGCCGAGTGGCTGGAAGACACACAGGAGGCC 2875
QY 629 hrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyG 649
Db 2876 AGTTCTGCACACACAGCCAGCGCTGCAGGCTATGAGCGGAGGAGCAGATGGCA 2935
QY 649 luGlu-----GlnLysProGluGlu- 655
Db 2936 ATGAGCTGGAAACGGCTCGGGCCGCGTGTATGGAGAGCCAGGCGGAGCAGCAGGAGG 2995
QY 655 ----- 655
Db 2996 GTGGCAGCAGGAAGGAGGTGGCGGCTGACCCAGGAGCGGGCGTCCCGAGCGTG 3055
QY 656 -----P 656
Db 3056 ACCTTGCCCTGGAGAGCGGCCAGACAGAGCTTGAGATGCGGCTGCAGAAACCCCTCA 3115
QY 656 roLysArgArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerL 676
Db 3116 ACAGCAGCGTGGAGTTCGCTACCTCCGCAAGAGGACACTGGCTCATGCCCTGACGAAA 3175
QY 676 yLysArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyG 696
Db 3176 AGGNA-----GGCAAGGACACAGAGTT-GGCCAAGCTTCGTGCTG 3216
QY 696 lyAspSer-----HisArgAlaGluGluAlaSerLysAspLysGluAla-----GlyT 712
Db 3217 GAGCAGCCAGATAAAAGAGCTGAGGAACTTCGGCAAAACCGTGAAGCAACTGAAGGAA 3276
QY 712 hrAspAlaValProAla-SerThrGlnGluGlnAspGlnAlaGlnGlySerSerPro 731
Db 3277 CA-----GCTGGCTAGAAAGAAAGAGCAGCAGCATCTGGCTCAGGAGCCAA 3324
QY 732 GluProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArg 751
Db 3325 TCTGAGCTCTGCGCAGCAGACAGCCAAACAGGCCCAAGCTG---GAAGCAGCTCGGGCA 3381
QY 752 LeuValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGlu----- 768
Db 3382 GAGGTGAGCAAGTGGAAACAGCAATGTCAGAGCAGCAGCAGGAGCAGGCTGACAGCTCGAA 3441
QY 769 -----AspSerSerValGluGlnLeu 775
Db 3442 CGCAGCTCGAGGCTGAGCGGCTCCCGGCTGAGCGGACAGTGTCTGAGACTCTG 3501
QY 776 SerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIlePro 795
Db 3502 CAGGCCAGTTAGGAGGAGAGGCCAGGAGCTAGGGCAGCATGAGTGCCTTAGCCCTCG 3561

QY 796 GlyArgArgLysLysArgAla-----AspGlyLysGlnGluGlnAlaThrValGluAsp 813
Db 3562 GCCCAAGGGAGTGGCTGCTCCCGCACCAGGTACAAGCACACACAGAGGTGAAGAT 3621
QY 814 SerGlyProValGluIle-----AsnGluAspAspProAsnValProAlaVal 829
Db 3622 GAGTGAAGCCCGAGGTGGCCCGGGCGGCAAGAGCTGAGAGGAAAAATAGCTCATC 3681
QY 830 ValProLeu---SerGluTyrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsn 848
Db 3682 AGCAGCTTGGAGGAGGAGGTGTCCATCTGAATGCCAGGTCTCTGGAGAGGAGGG- 3738
QY 849 ThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeu 868
Db 3739 ---GAGACCAAGAGTTGAAGCGGCTGGTATGCCCGAGTCAGAGAAGACCAAGCTG 3795
QY 869 ValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGlu 888
Db 3796 GAGGAGAGTGCCTCTGCTC-----AGGCAG 3822
QY 889 ArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGlu 908
Db 3823 AGACGCCAGCA----- 3834
QY 909 AlaMetProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProVal 928
Db 3835 -----ACAGTGCAGAGCTGCAGAACGCAGCTCTGCTCTCGGGAGG- 3876
QY 929 LeuThrClnThrLeuProGluGlyLysAspAlaHisAspAspMetValThrSerGlu- 947
Db 3877 ---AGGTGCAGAGCTCCGGGAGGAGGCTGAGAAACAGCGGTGGCTTCAGAGAAC 3930
QY 948 -----ValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArg 965
Db 3931 CTGGCGCAGGAGCTGACCTCACAGGCTGAGCTGCGGAGGAGCTGGGCCAAGAATTGAAG 3990
QY 966 Thr-----GluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSer 983
Db 3991 GCCTGGCAGGAGAAAGTTCTTCCAGAAAGAGCAGCGCTCTCCACCTGCGAGTC- 4044
QY 984 AlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGlu 1003
Db 4045 -----GAGCACACCCAGCACAGGCC 4065
QY 1004 ValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGln 1023
Db 4066 CTGGTGGTGGAGTGTGCCAGTAAGCACCTCTGCAGCAGGCTGCAGGCCGAGCAGGCC 4125
QY 1024 AlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArg 1043
Db 4126 GCTGCCGAGAAACGCCACCGTGAAGAGCTGGAG-----CAGAGCAAGAGGCC 4173
QY 1044 ThrGlySerLysAlaLeuGlu----- 1050
Db 4174 GCTGGGGAGCTGGCGGAGAGCTGCTGCCGGCCAGCGGAGCTTGGGGAGCTGATTCCT 4233
QY 1051 -----LysValGluGluValGluAspSerGluValLeuAlaSerGluLys- 1066
Db 4234 CTGGCGCAGAAAGTGGCAGAGCAGGACCGAACAGCTCAGCAGCTCGGGCAGAGAGGCC 4293
QY 1066 ----- 1066
Db 4294 AGCTATGCAGAGCAGCTGAGCATGCTGMAAGAGCGCATGGCTGCTGGCAGAGGAGAAC 4353
QY 1066 ----- 1066
Db 4354 CGGGGGCTGGGTGAGCGGGCCAACTTGGCCGCGAGTTTCTGGAAGTGGAGTTGGACCAG 4413
QY 1067 -----GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeu 1084
Db 4414 GCCCGGAAAGTATGTCACAGAGTTGGCAGCGCTGAGTGTGCTGATGCTGAGACCCCTG 4473
QY 1085 AlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThr 1104

```

Db 4474 GCTGAGGTGACGCGAAGACACAGACACTGCGCCGGAGCTGAGGTG-----ATGACT 4527
Qy 1105 AlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGlnGln 1124
Db 4528 GCCAAGTATGAGGGTGC-----AAGTCAAGGTCTCTGAGGAGAGG 4569
Qy 1125 AlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrPro 1144
Db 4570 CAGCGGTTCAGAGAAAGAGAGCAAACTCAGTGCAGGTGAAAGACTGATAGAAA 4629
Qy 1145 LeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAsp-----GluThrIleAspSer 1162
Db 4630 CTGGCTGACTCTGACCAACCCAGCAAGGTGCGACGACAGAACTGAGAGGTCTCCAGGCT 4689
Qy 1163 Gln-----AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGlnGlnGln 1180
Db 4690 CAGGAGGCGAAGAGCGAGAGGCGCGCTTCAGGCCAGCTGAATGAATGAACTGCA 4749
Qy 1181 AlaAlaThrAlaGlnLysGlnGluProSer-----ThrLeu 1192
Db 4750 GCCCAGTTGAGCGAGAGAGGAGCGAGCGCTGAGCACTATAGCTGCAGATGAGAAAGCC 4809
Qy 1193 ProAsnAsnValProAlaGlnGlnGluIleGlyGlnGlu-----ProGly 1207
Db 4810 AAAACACATTATGATGTCGCAAGAGGAGCGCAAAACCAAGACTGCGAGAGCACTCGCGAGC 4869
Qy 1208 ArgAspValLeuGlnProThrGlnGlnGlnLeuThrAlaAlaValProValLeuAla 1227
Db 4870 CTGGAGCGACACTGACAGAGAAACAAAGAGCTGCGAGCTGAAGCT----- 4914
Qy 1228 LysThrGlnValGlyGlnGlnGlnValAspTrpLeuAspGlyGlnLysValLysGln 1247
Db 4915 ---GAACGCTGCGGCGATGATACAG-----CAGCTGGGCTGAAGCAAGAGAG 4962
Qy 1248 GlnGlnGlnValPheValHis-----SerGlyProAsnSerGlnLysAlaAlaAspVal 1265
Db 4963 GCTGAACAGACCTGCCGCGACCTTACTGCGCGAGCTCGAGCGCTGAGGACAGAGTGGCC 5022
Qy 1266 ThrTyraAspSerGlnValMetGlnValAlaGlyCysGln-----GluLysGln 1281
Db 5023 CATGCAGACACAGAGCTTGAGACCTGGCAATTCAGAGTGGCAACTGATGCTTAAAG 5082
Qy 1282 SerThrGlnValGlnSer-----LeuSer 1289
Db 5083 ACCCGGAGACCCCGAGCTGAAGCCCGCAGCTGGAATTGACAGCCTGATCTGAGAC 5142
Qy 1290 LeuGlnGlnGlyGlnMetGlnThrAspValGlnLysGlnLysArgGlnThrLysProGln 1309
Db 5143 TCGAGAGAGGGAGCCCACTCACTATCACAGCAGAGCTGCTGT---ACCCAGCCAGAC 5199
Qy 1310 GlnValSerGlnGlnGlnGlnGlnGln-----ThrAlaAlaProGlnLysGlnGly 1326
Db 5200 GGCACACAGGTCCTCTGGAACACAGCTCACTATCTCCA-GCGCTCGCCCGCAAGGT 5258
Qy 1327 ThrTyrglyLysProValIleThrLeuAspMetProSerSerGlnLysArgGlyLysAlaLeu 1346
Db 5259 AGAATCCCTGAGAGAGTCTTCACTTCACTCCATCCCTGCTCGGAGTCAAGGCCCTCT-- 5315
Qy 1347 GlySerLeuGlyGlySerProSerLeuPro----- 1356
Db 5316 -----GAGAGAGCAGCTGAGCTCCCTGGGAGACGTCCTTCTGGACTCGGGTGTAA 5366
Qy 1357 -----AspGlnAspLys 1360
Db 5367 GACCCGCTCGCTCGTGGCGGACACAGAGATCATCATCATCATCATCATCATCATCATCAT 5423
Qy 1361 AlaGlyCysIleGlnValGlnValGlnSerLeu-AspThrThrValThrGlnThrAlaGln 1380
Db 5424 GCTAGATGTGAAGAGACGACAGCCCAACTCATCTATCTACAGACCGCGGTCTCTCC 5483
Qy 1380 ValAlaValGlnLysValIle-----GluThrValValIleSerGlnThrGlyGln 1396

```

```

Db 5484 TGCTTCCAGGCTAGCGCTGCGAGGACCACTCTCTTACTCAGTCTAGTCTGCGCTGGT-- 5541
Qy 1396 UserProGluCysValGlyAlaHisLeuLeu----- 1406
Db 5542 -TCTCCCATTTATGGCAACTCAGCCCTGCTCAGCTTGCCTGCTACCGCCACCACACTCG 5600
Qy 1407 ----ProAlaGlnLysSerSerAlaThrGlyGlyHisIleThrLeuGlnHisAlaGlnAs 1425
Db 5601 CAGTTCTCTGCTGCTTCCAGGCGCG-GGTGTCCAGTGGGGCCCTTCAGAGAAAGAAACA 5659
Qy 1425 PThrValPro-----LeuGlyProGluSe 1433
Db 5660 GCTTCTACATGGGACCTTGCCAGATGAGCTGAGCAGCTGATGATGATGATGATGATGATGAT 5719
Qy 1433 GlnAlaGlnSerIleProIleIleValThrProAlaProGluSerThrLeuHisProAs 1453
Db 5720 CAGACTCTCAGCAGGCGCAATCAGATGTGCCCCCAGCATCTGAAAGACCTGATATCCCTCG 5779
Qy 1453 PLeuGlnGlyGlnLysSerAlaSerGlnArgGlnArg---SerGlnGlnLysAspLysPr 1472
Db 5780 AGTCCAGGCTTCTCTGAGCTGGGACCATTCACAGATGAGAGAGTGAAGAACTGGAGAC 5839
Qy 1472 QAsp-----AlaGlyProAspAlaAspGly 1480
Db 5840 CCCAAGAGACCTTGGCGCGAGCCAGCATGACGACCATCAATCAATGACCGAGGG 5891

RESULT 13
US-08-931-999-4
; Sequence 4, Application US/08931999
; Patent No. 6043219
;
GENERAL INFORMATION:
;
APPLICANT: Iandolo, John J.
APPLICANT: Crupier, Scott S.
TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Hovey, Williams, Timmons & Collins
STREET: 2405 Grand Boulevard, Suite 400
CITY: Kansas City
STATE: Missouri
COUNTRY: U.S.A.
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,999
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/710,561
FILING DATE: 19-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25043-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6755 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: UT0007

```


Db 1771 AAAAAACAAAGCAGAAAAAGAAAGAAACAAACAAACAAACAAAGCACA 1770
 QY 810 rValGluAspSerGluProValGluIleAsnGluAspProAsnValProAlaValVa 830
 Db 1771 AAGAAAAAGCG-----GAGACAAACAAACAAACGCGGCAAGAC 1812
 QY 830 LProLeuSerGluTyrAsnAlaValGluArgGluAspMetGluAlaGlnGlyAsnThrGl 850
 Db 1813 CGGACACAGCAAA---CGAAGAAAAAGCGAAACAAAGAAAGAAAGAAAGAAAGAA 1869
 QY 850 uLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerTyrThrLeuValHt 870
 Db 1870 AAAACACACA-----GCCGACCGCAGAGAGAG----- 1897
 QY 870 sThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSe 890
 Db 1898 -----GAGACGAAAGAAAGCAAGCAAGCAAGAAAGAAACGCAAA 1935
 QY 890 rProSerTrpIleSerAlaSerValThrGluProLeuGluHtSThr-----AlaGlyGl 908
 Db 1936 A-----GAAAGAAAAACACGAGCCCGCAAGAAAGAAAGAAAGAGAGAGAGA 1983
 QY 908 uAlaMetProProValGluGluValThrGluAspIleIleAlaGly---GluThrPr 927
 Db 1984 AACACAGCCCGAGAAAAAGACAGAAAAACCGAGAAACAAAGAAAGCAGAGGAAAGGCC 2043
 QY 927 oValLeuThrGlnThrLeuProGluGlyAspAlaHisAspAspMetValThrSerGl 947
 Db 2044 CCCAACGAGGAAAGCAAGAAAGAAAGAAAGAAAGAGCCACCAACAAAGCAAGAAACAGA 2103
 QY 947 uValAspMetThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGl 967
 Db 2104 AGCGAGGGGGAAGAAAGCAAGAAAGAGGACACAGAAACGAAAGAAAGACCAAGAA 2163
 QY 967 uGluValThrGluAlaSerGly-----AlaGl 976
 Db 2164 GAGAGGAGCAAAACAGAAAAAGGGGAGAAACAAAGAAAGAAAGAAAGAAAGCCCA 2223
 QY 976 uGluThrThrAspMetValSerAlaValSerGluLeuThrAspSerProAspThrThrGl 996
 Db 2224 ACGAACCGAAAGCCGAAAGGCGCA-----ACGAGAGCAAGAAAGAAAGCGAG 2271
 QY 996 uGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGluArg 1016
 Db 2272 AAGACCAACAGAGG-AGAAAAAAAGACA-----GAAACCAAGAAAGAAAG 2318
 QY 1016 gGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValPr 1036
 Db 2319 GAAGACCCAAAGAAACAAACAAAGCAACGACAAACAAACAAACAAACAAACAAACAA 2361
 QY 1036 oAlaThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGl 1056
 Db 2362 -----ACAAGGCGCAACAAAGAAAGAAAGAAAGCAAGGAG----- 2403
 QY 1056 uGluAspSerGluValLeuAlaSerGlyLysGlyLysAspValMetProLysGlyProVa 1076
 Db 2404 -CAACCAAGCAAGCAACCAAGCAACCAACCAACCAAGCAACCAACCAACCAACCAAG 2461
 QY 1076 Lgln-----GluAlaGluAlaGluLysLeuAlaGlnGlySerGluThrGluGlnAlaTh 1094
 Db 2462 ACACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2518
 QY 1094 rProGluSerLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGl 1114
 Db 2519 AAACGAAACCCCGAAACCAAGCGCCCGCAGAGAAACAAAGAAAGCAACCAAGCAAGCA 2578
 QY 1114 nValIleLysLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSerGluThrIle 1134
 Db 2579 A-----GAAAGGCCACAGAGCGCACAGAGCGCACAGAGAGGCGCAAG 2614
 QY 1134 uThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyTh 1154
 Db 2615 GAGAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGAG 2674

QY 1154 rGlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValArg----- 1172
 Db 2675 ACAGAAAGAAAGAGCGCAAGCAACCAAGCAGACAAACAAAGAAAGAGGGAACAGAAACGA 2734
 QY 1173 -----GlnSerGlnValThrGluGluGluAlaAlaThrAl 1184
 Db 2735 CAAGAGAGGCAACAGACAAACAAACAAAGCAAGCAAGCAAGCAAGCAAGCAAGGCGGC 2794
 QY 1184 gLysGluGluProSerThrIleProAsnAsnValProAlaGlnGluGluHtSGlyGl 1204
 Db 2795 AAAGAACCGCCAAACGCCACGCAACCAACCAACCAAGGGGAAAGACGAGCGGAGC 2854
 QY 1204 uGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValPr 1224
 Db 2855 AGACGCGCGCAGGAC-----AAACCAACCAACCAACCAAGGAGCGGCAAC 2899
 QY 1224 oValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLy 1244
 Db 2900 GAAAGAAACAGAAAGAAACCCGAGAGAAACGAAAGCAAGCAAGCAAGCAAGAGAG 2959
 QY 1244 sValLysGluGluGlnGluValAlaPheValHisSerGlyProAsnSerGlnLysAlaIAs 1264
 Db 2960 AGAAAGAAAGAGACAGACACACAGAAAGCCACAGAAACAAACAAAGAAAGAGGGA 3019
 QY 1264 pValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGl 1284
 Db 3020 CACAAAGCGGGCAGC-----GGGACAGAAACCAACAGAAACCAAGCAAGCAAG 3061
 QY 1284 uValGlnSerLeuSerLeuGluGluGluGluMetGluThrAspValGluLysGluLysAr 1304
 Db 3062 ACCGAAACCGAGGAGCAAGAAAGGGGAGAGAA-ACCAACAAACAGAAAGAAAGACAAAG 3120
 QY 1304 gGluThrLysProGluGlnValSerGluGluGlyGlnGlnGluThrAlaAlaProGluHt 1324
 Db 3121 ACA-----CCACAGAAAGCAACCCAAAGCAAGCAAGAAAGAGGGAACCAAGAA 3174
 QY 1324 sGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluArg--Gl 1343
 Db 3175 ACGAAACCAACCGCCCA-----CAGCCAAACAAAGCCGACAGG 3216
 QY 1343 LysAla-----LeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAl 1361
 Db 3217 AAGACAGAGCAAGAAAGGCGCAAGCAGAGGCAAC-----GACAA-- 3256
 QY 1361 aGlyCysIleGluValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAl 1381
 Db 3256 ----- 3256
 QY 1381 aValGluLysValIleGluThrValValIleSerGluThrGluGluSerProGluCysVa 1401
 Db 3257 -----GAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3277
 QY 1401 LglnAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHtSTrThrLeuGl 1421
 Db 3278 -----CCCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3311
 QY 1421 nHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleGl 1441
 Db 3312 GCACAGAGAGCAACAGCAAGAAAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3360
 QY 1441 eValThrProAlaProGluSerThrLeuHisProAsp-----LeuGlnGlyGluIleSe 1459
 Db 3361 -----GAAAGCAAGAAAGAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3404
 QY 1459 rAlaSerGlnArgGluArgSerGluGluGluLysProAspAlaGlyProAspAlaAs 1479
 Db 3405 GAAGAAAGAAAGAGAAAGAAAGCAAGCGAAAGAAAGAGGAGGAGAGAAAGAAAGAA 3464
 QY 1479 pGlyLysGluSerThrAlaIleGluValLeuLysAlaGluProGluIleLeuGluLe 1499
 Db 3465 GAACAAAGAAACCGAGCAAGAGAAAGAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAG 3524

Alignment Scores:			
Pred. No.:	3.23e-17	Length:	11236
Score:	388.00	Matches:	340
Percent Similarity:	37.54%	Conservative:	261
Best Local Similarity:	21.24%	Mismatches:	588
Query Match:	4.81%	Indels:	416
DB:	1	Gaps:	82
US-09-302-432-4 (1-1596) x US-07-853-913-1 (1-11236)			
Qy	13	GlucInProAlaGlySerAspThr-----ProSerGluLeuVal-----	25
Db	6721	GAGAACTCTGGGGCAGCAGCAGGCTCACTTCCTGTGATGATCTACCTCTTACCCACA	6780
Qy	26	--LeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAsp	44
Db	6781	AACCTCAACCTCACCACTTCTTAGAGGCTAAGAT-----GGAGATCCAGTGAG	6834
Qy	45	AlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAsnGlyVal	64
Db	6835	TCTAGAGTTTCTAGCATATTCCAGGAAGATAGGGGCAATCTGGGAACCTG-----GTA	6888
Qy	65	AlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGluGlyGln-----GluGlu	82
Db	6889	GAGAAAGAACAGATAGAGGTAAAGTAGAAACACGCTCAGCCCAACAAACACAGAA	6948
Qy	83	GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg	102
Db	6949	AGTGCTCTGACACAGAGAACCACG---GATTCAGGACCTTCAGAAAGAAACA	7005
Qy	103	ValGluGluMetAlaAlaAsn-----SerThrAlaValGluAspIleThrLysAsp	119
Db	7006	CTGAAGCTCTAGAGAGGACCACTGATGTCTCTGAAAATCCAGAACTATGACACGA	7065
Qy	120	GlyGlnGlu-----GluThrSerGlu	126
Db	7066	GGGAAAGAGAATTGCAATTTCTTACAGAAAGCCACCTGGGAACACTAGAAGGCCAGAA	7125
Qy	127	IleIleGluGlnIlePro-----AlaSerGluAsnAsnValGluGluMetValGln	143
Db	7126	AAAGAAAGCAATACCATAAGTCTTTAGAAAGAAAGATGTAGAG-----	7173
Qy	144	ProAlaGluSerGlnAlaAsnAspValGlyPheLysValPheLysPheValGlyPhe	163
Db	7174	-----TCAGAGAAAACCTCTAGAAAATGGGTTCTGTACTCTGAGCTTTTAGGAAA	7227
Qy	164	LysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuThrVal	183
Db	7228	GAAGACACAAAGACAGAGGATCAAGAAATTAATGTCTCTCAAGGT-----ACACTA	7278
Qy	184	LysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluPro	203
Db	7279	AAGAGATTTTCATCTCTAGGAAAGGAAGTCAAGAAGTAGTGAGGCTTTCAAAAGAGGGG	7338
Qy	204	SerValGluThr-----AlaValGlyGluSerAla	213
Db	7339	AACCTAGAATCATGGACAGCTTTAAAGAGGAGAGCCCAACCCACTGGATTTCCAGGA	7398
Qy	214	SerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLeu-----	230
Db	7399	GCTGAGCACCATGTCTTGAGACAGCTGGTAGAAGAGAGGATCAGAGCTTCCCAAGGTCT	7458
Qy	231	--LysGlnGluGlnSerSerThrGluIleProLeuGlnAlaGluSerAspGln---Ala	248
Db	7459	CCAGAGGAAGAGGACCGAGGCGATGTAGACCTCTGCAGAAAGAGAATCAGGAACCCACTA	7518
Qy	249	AlaGluGluAlaLysAspGluGlyGlu-----GluLysGlnGlu	262
Db	7519	GGGTATGAAGAACA-----GAGGCCAGATACTTGAGAGACTCATATGAAAAGAGAGT	7572
Qy	263	LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrSer	282
Db	7573	CAGAGTCTCTGAGGTCTCCAGGAAGAG-----GACCAGAGCGCAGGTAGA	7620

QY 283 SerPheLysPhePheThrHisGlyTrrPalaGlyTrrArgLysLysThrSerPheLys 302
 |||
 |||
 Db 7621 TCTCGCAAGAAA-----GAGAAATCAGAGCCACTAGGATATGAA 7659
 QY 303 LysSerLysGluAspAspLeuGluThrLalGluLysArgLysGluGluGluLysLys 322
 |||
 |||
 Db 7660 GAAGCAGAGGACCATGCTTGAGAGCATGATAGAAAAAGAGCTGAGATCCCTGAAG 7719
 QY 323 ValAspGluGluGluLysGluLysThrGluProLalAspSerGluGluGluLysLys 339
 |||
 |||
 Db 7720 TCTCCAGAAAGAAACCATGGAAGCCCTTAGAAAGAGAAATCAGAAATCTCTG 7779
 QY 340 -----ProLalGluLysThrAspGluLysLys 348
 |||
 |||
 Db 7780 AGCATCTTGAAGAAACCAGAGACTTTGTATCCATGAGAAACAGAGCAAGAGCCCA 7839
 QY 349 LeuSerAla-----AspTyrGluLysValGluLeuProLeuGlu----- 361
 |||
 |||
 Db 7840 CTGAGATCTCTAGAAAGTAGAAGAGGAGAGCAGAGAAATTGTGAACCTCTAGAAAAAGTG 7899
 QY 362 -----AspGluValGlyAspLeuGluLalAspSerGluGluLysCysAlaProLeuAla 379
 |||
 |||
 Db 7900 ACTCAGGATTCCTCGACTCTCTA-----GCAGAAAGAGATGTGACGCCACTGAGG 7950
 QY 380 ThrGluValPheAspGluLysMetGluLalHisGluGluValAlaGluLysValHisVal 399
 |||
 |||
 Db 7951 -----TATCTGGAAAGAGATGATGCTCATTAATTAAGAGCCTTCTAGAA----- 7992
 QY 400 SerThrValGluLysThrGluGluGluGluGlyGlyGlyGlyGluLalGluGluGlyVal 419
 |||
 |||
 Db 7993 -----GACAAGACTCACAAGATGCTCTTGGGCTCTCTGGAAGATGAAATGGGATAC 8043
 QY 420 ValVal-----GluGlyThrGlyGluSerLeu-----ProProGluLys----- 432
 |||
 |||
 Db 8044 ATTATTATATCCACAGAAAGTAGAGCCAGGTTTCAATTGAGCCCTCAGAAAGAGAGAGAC 8103
 QY 433 -----LeuAlaGluProGluGluValProGluGluLalGluPro 445
 |||
 |||
 Db 8104 CAGAGATTGTGAACCATCTAGAAAAAGAAAGTCAAGAGTCTCGAGCTTCTCAGAGAA 8163
 QY 446 AlaGluGluLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGluLeu 465
 |||
 |||
 Db 8164 GAAGACCATGGTATG-----GAGAGATCTCTAGAAAGAGAGAAACCATGATCATCTG 8214
 QY 466 ThrAspLysSerProGlu-----GluLysThrLeuProLysHisProGluGly 481
 |||
 |||
 Db 8215 AGTTCTGTGAAAAAAGAGAGACCATGATGTTGAGAGCCCACTAGAAAGAGAGTCAAGAC 8274
 QY 482 LLevalSerGluValGluMetLeuSerSerGluGluLysGlyLeuValGluGlySerPro 501
 |||
 |||
 Db 8275 TCAGGAGATCTCTTGAAAGT--GAGAGCCAGAGAGACCTTT-----GGACCT 8319
 QY 502 LeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysLysGlyLysGly 521
 |||
 |||
 Db 8320 CTGGAAAAA---GAGATGAGAGTCCCTGAGATCTCTACAGAGACAGACCAA----- 8370
 QY 522 LysArgGlyGlyGlyGlyAspGluGluProGlyGlyLysLysLysLysLysLysLys 541
 |||
 |||
 Db 8371 -----GAGGAACAGAAAGCTTGAACAGAGAC 8397
 QY 542 ProGluSerAlaAspGluGluLysGlyGluSerSerAlaSerSerProGluGlu----- 559
 |||
 |||
 Db 8398 CAACAAACACTGAGGCTGAGGAATGAGCAGATGCGACTGAGCCACAGAAAAAGGTG 8457
 QY 560 ---ProGluGluThrThrCysLeuGluLysGlyProLeuGluLalProGluAspGlyGlu 578
 |||
 |||
 Db 8458 GATCCAGAG-----TTACCGAAG--CTCTT-----GGAAT 8487
 QY 579 AlaGluGlyGlyThrThrSerAspGlyGlyLysArgGlyGlyLysLysLysLysLysLys 598
 |||
 |||
 Db 8488 GACCAAGAAATAGCTATCTCTTGAAAAAGAAATCAAGAGTCACTAGTG----- 8538

QY 599 SerPheLysLysMetValThrProLysLysArgValAlaArgArgProSerGluSerAspLys 618
 |||
 |||
 Db 8539 -----TCACTGAAA 8547
 QY 619 GluGluGluLeuGluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSer 638
 |||
 |||
 Db 8548 GAAAAAGTATAGAGACATGATGATGATCTTTAGAA-----ACA 8583
 QY 639 GluMetGluAspGluValLysThrValGlyGluGluGluLysProGluGluProLysArg 658
 |||
 |||
 Db 8584 GAGATCATAGAAACCATCTGAGACATGCAAGAGAGCCTG-----GAAAGAGAGAG 8634
 QY 659 ArgValAspThrSerValSer---TrrGluLalLeuLysLysValGlySerSerLysLys 677
 |||
 |||
 Db 8635 TCTATGATCTCAGAGACCATTTGTG-----TCTACTGAA 8670
 QY 678 ArgAlaArgLysAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyGly--- 696
 |||
 |||
 Db 8671 GTGCTTAGAGAGACATGAAACCTCCAGAAAGATAGCCCCCAGAGATCCCTAGGCTCTGTG 8730
 QY 697 AspSerHisArgAlaGluGluLalAspLysAspLysGluAlaGlyThrAspAlaValPro 716
 |||
 |||
 Db 8721 GATGAGAACCCAGAGACATGACATCCCTGAAAAAGAGAGTCA--AGAACTGAGCTCTCT 8789
 QY 717 AlaserThrGluGluAspGluLalGluLysSerSerSerProGluProLalGlySer 736
 |||
 |||
 Db 8790 GGGCAAGTGCAGAGACGTAAGACCGGTAGAGACATGACGA---GTGCTGCAAGTACA 8846
 QY 727 ProSerGluGlyGluGlyValSerThrTrrPgluSerPheLysArgLeu-----ValThr 754
 |||
 |||
 Db 8847 AGAGGCTGCAGAGAGAAC--AGCACAA--GAGTCTGAGAGAGCTGAACAGAGAGAGCTG 8904
 QY 755 ProArgLysLysSerLysSerLysLeuGluGlyValAlaGluAspSerSerValGluGlu 774
 |||
 |||
 Db 8905 CCTACTCTGGAATTAACAAGCGGTGGAGATGTGTGGAGGGCAAGACATGGGTGAG 8964
 QY 775 -----LeuSerThrGluLeuGluProSerArgGluGluGluSerTrrValSerLysLys 791
 |||
 |||
 Db 8965 GAGCACTCTGGCAACACAGAGAGTGGAACTGAGATTAAGGACAGATGATGTG--- 9021
 QY 792 LysPheLeuProGlyArgArgLysLysArgAlaAspGlyLysGluGluLalThrVal 811
 |||
 |||
 Db 9022 -----AGGGGCGCAAGTGTGAGAGGAAGAAAGCTCCACCA 9054
 QY 812 GluAspSerGlyProValGluLeuAsnGluAsp----- 822
 |||
 |||
 Db 9055 GAGGGA-----GAGCTGTTCAGAGATATTGTGGGGAGAGCCTGAGTCTGGG 9102
 QY 823 -----AspProAsnValProAla----- 828
 |||
 |||
 Db 9103 AGCTTGAGCCCAAGAGACAGAGGTCCTGCTGAGGCCCTCGACACCTGGAAGAGAGG 9162
 QY 829 -----ValValProLeuSerGluTyr-----AsnAlaValGluArgLysLysMetGlu 844
 |||
 |||
 Db 9163 GCCTTAGAGTCCAGTGTGCTAGTCATGCCAGAGGTGACAGACGAGATGAGATAGA 9222
 QY 845 AlaGluGlyAsnThrGluLeuProGluLeuGlyAlaValTyrValSerGluGluLeu 864
 |||
 |||
 Db 9223 GCCCAAGCAGGTGAACA-----GACTCCATA 9249
 QY 865 SerLysThrLeuValHisThrValSerValAlaValLLeuAspGlyThrArgAlaValThr 884
 |||
 |||
 Db 9250 GAGGTGACCTT-----GGTTAGAGGCTGCCAGAACTGAGCTG 9288
 QY 885 SerValGluGluArgSerProSerTrrPheSerAlaSerValThrGluProLeuGluHis 904
 |||
 |||
 Db 9289 GAATCGACAGAGAA-----GTGATAGGCTTAGAGACCAAGGATTTT 9333
 QY 905 ThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspLysLysLysLys 924
 |||
 |||
 Db 9334 GCCAGGAGAGAGCCATTCCTCCA-----TCCCTGGGGAGAGAAAGTGTGAAGCAAG 9387
 QY 925 GluThrProValLeuThrGluThrLeu-----ProGluGlyLysAsp 938

QY 533 luTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSerS 553
DB 4223 TCAGAGCAGCACTGAACCGGAGAGAGTGGAGACTTCGAGCAGGAGGAGCGCGACG 4282
QY 553 erAlaSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluA 573
DB 4283 AGCGGCTGAACCGCGAGAGCGCGGAGAGAGAGCGGCCACGAGCTCTGAGAGCGAGG 4342
QY 573 laProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysArgGluG 593
DB 4343 AGCAGGAGGAGCGCGCCACGAGCACTGAGCGCGAGCAGCAGCAAGGCGCGAGCAGC 4402
QY 593 lyIleThrProTrpAlaSer-----PheLysLysMetValThrP 606
DB 4403 GGCTGAACCGCGAGGAGGAGAGAGAGGCTCGAGCAGCGGCTGAAGCGCGCATGAGG 4462
QY 606 roLysLysArgValArgProSerGluSerAspLysGluGluGluLeuLysValL 626
DB 4463 AAGAGAGCGCGAGCAGCAGCTAGCTGAGGAGGAGCAGGAGCGCGCGGAGCGGATTA 4522
QY 626 ysSer-----AlaThrLeuSerSerThrAspSerThrValSerG 639
DB 4523 AGAGCGCATCCCGAGTGGCAGTGGCAGCTAGAAAGCGAGGCGCGCAGCGCAAGCA 4582
QY 639 luMetGlnAspGluValLysThrValGlyGluGluGlnLysPro-----GluGluProL 657
DB 4583 AAGCTTTACTCGAGGCGCGCGCAAGCAGAGAGGCGGAGCGCGCGCAAGAGCAGGAGGAAA 4642
QY 657 ysArgArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysL 677
DB 4643 AGAGCGCGCGGAGGAGTGCAGTTCGATGGGCG-----GAGGAGG 4681
QY 677 ysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArgThrLeuGlyGlyA 697
DB 4682 AACGGGCTCACCGGCAGCAGCAGGAGGAGGAGCGCGCGGAGCTTCACATGG----- 4734
QY 697 spSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProA 717
DB 4735 --CAGTGGCGAGCGCGGAGGAGAGG-----GAGAGGCGCGTCAGAGGCTGTGCG 4783
QY 717 laSerThrGlnGlnAspGln-AlaGlnClySerSerProGluProAlaGlySer 736
DB 4784 CAGGCGCCCATTCGGGAGCAGCGGAGAGGCGAGCTGAGGCGCGGAGCGCGCAGCAGC 4843
QY 737 ProSerGluGly----- 740
DB 4844 GGGACACACGGTTCTCCCGGAGGAGGAGGAGGAGGAGCGCGCGCGCGCGCGCGCG 4903
QY 741 -----GluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLys 757
DB 4904 AGAGGGAGAAAGAGCTGCAGTTCCTGGAGG-----AAGAGGAGCAGCTCCAG-CGGCGG 4956
QY 758 LysSerLysSerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLeuSerThr 777
DB 4957 GAGCGTCCCAACAGCTCCAGGAGGAGGAGGAGCGCGCTCCAGGAGGATCAGGAGGAGG 5016
QY 778 GluIleGluProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArg 797
DB 5017 CGACAGGAGCAGCGCGCGGACCAAAATGG-----AGGTGGCACTAGAGAGAGAA 5067
QY 798 ArgLysLysArg-----AlaAspGlyLysGlnGluGlnAlaThrVal 811
DB 5068 AGGAAGAGAGCGCGCCACACGCTGTACGCCAAGCAGCGCTTACAAGAGCAGCTCAGGAAG 5127
QY 812 GluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAlaValPro 831
DB 5128 GAACAGCAG-----CTGCTCAGGAGGAGGAG----- 5154
QY 832 LeuSerGluTyrAsnAlaValGluArgGluLysMet----- 843
DB 5155 ---GAGGAGCTACAGAGAGGAGCGCGGAGAGAGAGGCGCGCAAGAGACAGGAGAGCA 5211

QY 844 -----GluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVal 860
DB 5212 TACCGCGAGGAGAGCAGCTGCAGCAGGAGGAGAGAGCAGCTGCTG----- 5256
QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
DB 5257 AGAGGAGAACCGGAGGAGAGAG-ACGCCAGGAGCGGAGGAGCAATATCGGAAGGATAA 5315
QY 881 ArgAlaValThr-----SerValGluGluArgSerPro 891
DB 5316 GAAGCTCGCAGCAGAGAGAGCAGCTGCTGGGAGAGAACCGGAGAGAGAGAGAGCGCCA 5375
QY 892 SerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro 911
DB 5376 GGA-----GCGGAG 5396
QY 912 ProValGluGluVal-----ThrGluLysAspIleIleAlaGlu 924
DB 5397 CGAGGAGAGAGAGCTTCAGCAGGAGGAGAGAGCAGCTGCTGAGAGAGAGAACCGGAGAGAG 5456
QY 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
DB 5457 AAGCGCGCAGGAGTGGGAGAGGAGTACCGCAAAAGACGAGCT----- 5501
QY 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThr-GluThrSerGluAlaLe 964
DB 5502 -----GCAGCAG 5525
QY 964 uArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAl 984
DB 5526 GAGAGAGAGAACGGGAGAGAGAG----- 5547
QY 984 aValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVa 1004
DB 5548 -----AGACTCCAGGCGCGGAGAGGCAATATCGGAGAGAGAGAGAGAGAGAGAGAG 5600
QY 1004 lGluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIle----- 1021
DB 5601 GGAAGAGCAGCTGCTGGGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5660
QY 1022 -LeuGlnAlaValAlaAspLysValLysGluGluSerGlnVal-----ProAl 1037
DB 5661 ATATCGAAGAGAGAGAGAGCTGCAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5720
QY 1037 aThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluG 1057
DB 5721 GAAGAGAGCGCGCAGGAGCGGAGAGGCAATGTCGGAGAGAGAGAGAGAGAGAGAGAGAG 5780
QY 1057 uAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValG 1077
DB 5781 GGAAGAGCAGCTGCTGAG 5825
QY 1077 nGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSe 1097
DB 5826 -----GGAGCTGGAGAGCAATATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5878
QY 1097 rLeuGluValProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLy 1117
DB 5879 AACAGCGATACCGGATGAGGATCA----- 5903
QY 1117 sLeuGlnGlnLeuMetGluGlnAlaValAla-ProGluSerSerGluThrLeuThrAspS 1137
DB 5904 -----GGCAGTGTATCTGAATGCGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 5959
QY 1137 er-----GluThrAsnGlySerThrProLeuAlaAspSerAspT 1150
DB 5960 ACAAGGTTTACTCAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6019
QY 1150 hrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaA 1170
DB 6020 TGGCGGAG 6061
QY 1170 laValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProS 1190

```

Db 6062 GTGAACAGCAGAGAGAGATTCCTGACACAGAGAGAGCGCTGCGACGCGCAACAGGC 6121
      ::|||:::|||||:::
Qy 1190 eRThLeuProAsnAsnValProAlaGlnGluGlnHisGlyGluGluProGlyArgAspV 1210
      ::|||:::|||||:::
Db 6122 ATTTCCAGAGAGAAACAGCTCGAGCGAGAGAGCAAAAGAGAGCAAAAGCGCGACA 6181
      ::|||:::|||||:::
Qy 1210 alLeuGluProThrglnGlnGlnLeuThrAlaAlaValProValLeuAlaValThrg 1230
      ::|||:::|||||:::
Db 6182 GGAACTCCCAAGAGAGAAAGCAGTTGCTGAGAGAG-----GAAAGAGAG 6226
      ::|||:::|||||:::
Qy 1230 luValGlyGlnGlnGluValAspTrpLeuAspGlyGluValLysGlnGlnGlnG 1250
      ::|||:::|||||:::
Db 6227 AGAAGAGAGCGCTCAAGAGAGAGAGC-----AGAAATTCCTCCGAGAGAGAAC 6274
      ::|||:::|||||:::
Qy 1250 luValPheValHisSerGlyProAsnSerGlnLysAlaAlaAspValThrTyrAspSerG 1270
      ::|||:::|||||:::
Db 6275 AGCTGCTC-----CAGAAAGAGAGAGAGAGCGCGCTCCGCGCAAG 6316
      ::|||:::|||||:::
Qy 1270 luValMetGlyValAlaGlyCysGlnGlnGluSerThrgluValGlnSerLeuSerL 1290
      ::|||:::|||||:::
Db 6317 AG-----CGTACAGAAATTCGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6355
      ::|||:::|||||:::
Qy 1290 euGluGluGly-----GluMetGluThrAspValGluLysGluLysA 1304
      ::|||:::|||||:::
Db 6356 AGGAAACAAGGGGAGAAATTCCTCGAGAGAGAGAGAGCGCTCGCGAGAGAGAGAGAG 6415
      ::|||:::|||||:::
Qy 1304 rgGluThrLysProGluGlnVal-----SerGluGlnGlnGlnGlnGlnGlnGlnG 1322
      ::|||:::|||||:::
Db 6416 AATTCCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6474
      ::|||:::|||||:::
Qy 1322 roGluHisGlu-----GlyThrTyrGlyLysProValLeu----- 1333
      ::|||:::|||||:::
Db 6475 CGCGACAGAAATTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6534
      ::|||:::|||||:::
Qy 1334 -----ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerL 1349
      ::|||:::|||||:::
Db 6535 TTCCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6590
      ::|||:::|||||:::
Qy 1349 euGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysLleGluValGlnValG 1369
      ::|||:::|||||:::
Db 6591 --GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6630
      ::|||:::|||||:::
Qy 1369 InSerLeuAspThrThrValThr-----GlnThrAlaGluAlaValGluLysValI 1386
      ::|||:::|||||:::
Db 6631 CGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6690
      ::|||:::|||||:::
Qy 1386 leGluThrValAlaLieserGluThrGlyGluSerProGluCysValGlyAlaHisLeuL 1406
      ::|||:::|||||:::
Db 6691 GACAGAAATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6728
      ::|||:::|||||:::
Qy 1406 euProAla-----GluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHis--A 1423
      ::|||:::|||||:::
Db 6729 --CCAAGAGCGTACAGAAATTCGCGCAACAGAG-----ACTGCGAGTCAAGAA 6777
      ::|||:::|||||:::
Qy 1423 laGluAspThrValProLeuGly-----ProGluSerGlnAlaGluSerI 1438
      ::|||:::|||||:::
Db 6778 CCAGAGAGAAATTCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 6837
      ::|||:::|||||:::
Qy 1438 lePro-----IleIleValThrProAlaProGluSerThrLeu-HisProAsp 1453
      ::|||:::|||||:::
Db 6838 TTCCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6897
      ::|||:::|||||:::
Qy 1454 leuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluLysLysProAsp 1473
      ::|||:::|||||:::
Db 6898 CGTGACAGAAATTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6957
      ::|||:::|||||:::
Qy 1474 AlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeuLysAlaGlu 1493
      ::|||:::|||||:::
Db 6958 AGCGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7017
      ::|||:::|||||:::
Qy 1494 ProGluIleLeuGluLeuGluSerLysSerAsnLysIleValLeuAsnValIleGlnThr 1513
      ::|||:::|||||:::

```

```

Db 7018 CAAGAGAGAAATTCATGAGAGAGAGAG----- 7044
      ::|||:::|||||:::
Qy 1514 AlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHisAlaTyrAspSerGln 1533
      ::|||:::|||||:::
Db 7045 -----CAGCAGCTGCGCGCCAGAGAGGCGCAACACAGAGAGAGAGAGAGAGAGAA 7098
      ::|||:::|||||:::
Qy 1534 ThrGlnValProAlaCysAlaArgLeuAspSerLysGluProAsnArgCysTrpThrLysMet 1553
      ::|||:::|||||:::
Db 7099 TTCGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7158
      ::|||:::|||||:::
Qy 1554 LysAspAlaLysMetLysHisProValProGlnProArgGluAspLeuGlnValLeuThr 1573
      ::|||:::|||||:::
Db 7159 CGTGACAGAAATTCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7217
      ::|||:::|||||:::
Qy 1574 ValLeuGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaPro 1593
      ::|||:::|||||:::
Db 7218 -----GCTGCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7268
      ::|||:::|||||:::

```

Search completed: December 13, 2002, 04:17:51
 Job time : 438 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 03:13:21 ; Search time 944 Seconds
(without alignments)
668.597 Million cell updates/sec

Title: US-09-902-432-4

Perfect score: 8073

Sequence: 1 MGNSTGTEQSPQAGSDT.....AWAQRKCLPRLQKAPVSK 1596

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09902432/runat_12122002_141453_25918/app.query.fasta_1.1735
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09902432@cgn 1.1.85 @runat_12122002_141453_25918
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8073	100.0	6160	9 US-09-902-432-3	Sequence 3, Appli
2	6005	74.4	5134	9 US-09-902-432-1	Sequence 1, Appli
3	4300.5	53.3	5346	10 US-09-738-877-2	Sequence 2, Appli
4	4300.5	53.3	6608	10 US-09-738-877-1	Sequence 1, Appli

5	4300.5	53.3	6608	10	US-09-880-107-3439	Sequence 3439, Ap
6	462	5.7	5361	9	US-09-742-096-2	Sequence 2, Appli
7	462	5.7	6152	9	US-09-742-096-1	Sequence 1, Appli
8	431.5	5.3	7997	10	US-09-864-761-17726	Sequence 17726, A
9	427	5.3	15231	10	US-09-917-800A-1505	Sequence 1505, Ap
10	418	5.2	14800	10	US-09-954-456-1501	Sequence 1501, Ap
11	411	5.1	7104	10	US-09-815-242-4580	Sequence 4580, Ap
12	411	5.1	7107	10	US-09-815-242-8291	Sequence 8291, Ap
13	408	5.1	10300	10	US-09-960-253-145	Sequence 145, App
14	398.5	4.9	4840	10	US-09-880-107-3423	Sequence 3423, Ap
15	395	4.9	11185	10	US-09-880-107-3311	Sequence 3311, Ap
16	388	4.8	3953	10	US-09-864-761-19041	Sequence 19041, A
17	382.5	4.7	7639	9	US-09-971-536-28	Sequence 28, Appli
18	382.5	4.7	7792	12	US-10-044-090-359	Sequence 359, App
19	379	4.7	7497	10	US-09-960-253-175	Sequence 175, App
20	378.5	4.7	10432	10	US-09-919-172-97	Sequence 97, Appli
21	377	4.7	7434	10	US-09-815-242-4761	Sequence 4761, Ap
22	377	4.7	7437	10	US-09-815-242-8869	Sequence 8869, Ap
23	371	4.6	6940	10	US-09-880-107-1724	Sequence 1724, Ap
24	365.5	4.5	5331	10	US-09-764-176-6	Sequence 6, Appli
25	355	4.4	4854	9	US-09-963-875-2	Sequence 2, Appli
26	351.5	4.4	1855	10	US-09-864-761-19708	Sequence 19708, A
27	346.5	4.3	4372	10	US-09-864-864-323	Sequence 323, App
28	346.5	4.3	4372	10	US-09-864-864-337	Sequence 337, App
29	345	4.3	10825	10	US-09-727-384-5	Sequence 5, Appli
30	342.5	4.2	7065	9	US-09-991-496-115	Sequence 115, App
31	342.5	4.2	7065	10	US-09-874-923-115	Sequence 115, App
32	341	4.2	3388	9	US-09-954-531-988	Sequence 988, App
33	341	4.2	3388	9	US-09-954-531-1382	Sequence 1382, Ap
34	341	4.2	3388	10	US-09-954-456-1602	Sequence 1602, Ap
35	341	4.2	3388	10	US-09-967-768A-245	Sequence 245, App
36	335.5	4.2	4780	10	US-09-962-436-287	Sequence 287, App
37	335.5	4.2	7596	10	US-09-954-456-2215	Sequence 2215, Ap
38	335	4.1	1958	10	US-09-864-761-2927	Sequence 2927, Ap
39	334.5	4.1	7035	10	US-09-815-242-8615	Sequence 8615, Ap
40	329.5	4.1	6114	9	US-09-938-842A-2700	Sequence 2700, Ap
41	327	4.1	6457	10	US-09-880-107-3389	Sequence 3389, Ap
42	323.5	4.0	5721	10	US-09-785-770A-15	Sequence 15, Appli
43	323.5	4.0	8121	10	US-09-785-770A-14	Sequence 14, Appli
44	323	4.0	7869	10	US-09-954-456-1921	Sequence 1921, Ap
45	316	3.9	4747	9	US-09-854-133-729	Sequence 729, App

ALIGNMENTS

RESULT 1
US-09-902-432-3
; Sequence 3, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6160
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-3

Alignment Scores: 0
Pred. No.: 8073.00
Score: 1596
Length: 6160
Matches: 1596

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-902-432-4 (1-1596) x US-09-902-432-3 (1-6160)
QY 1 MetGlyAlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThr 20
Db ATGGCGCGGAGGAGTTCCTCCAGCAGCGAGCCCGGACCGCGCGGAGCGACAG 92
QY 21 ProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyValAlaGly 40
Db CCGAGCGAGCTGGTGGTCTCAGTGCCATGGGCCCGACGCTGAGACCTCCGGAGAGCTGGA 152
QY 41 AspProAlaAspAlaAspProAlaThrIlyLeuProGluIlyAsnGlyGlnLeuSerSer 60
Db GACCCCGCGGAGCGGAGCCCGGACCAAGCTCCACAGAGAATGGCCAGCTGTCTTCT 212
QY 61 ValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGlyGln 80
Db GTCAACGCGCTGAGTGAACAGAGAGATGTCATGTCCAAGAGAAAACAGAGGGGCGAG 272
QY 81 GlnGluGluValValAspGluAspValGlyIlyArgGluSerGluAspValArgGluIly 100
Db GAGGAAGAAGTGTGTATAGATGTTGACAGCGAGCTCAGAGAATGAGAGAAAAA 332
QY 101 AspArgValAlaGluIlyMetAlaAlaAsnSerThrAlaValAlaAspIlyeThrIlyAspGly 120
Db GACCGAGTGAAGAATGGCGGCACTCCACAGCTGTTGAGATATCAAAAGATGGG 392
QY 121 GlnGluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnValGluGlu 140
Db CAGGAGGAGACATCAAGAAATTAATTGAACAGATCCCTGCTTCAGAAAAACAATGGAGAA 452
QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheIlySerValPheIlyPhe 160
Db ATGGTACAGCTGTGAGTCCAGGCTTAATGATGTGGCTTCAGAAAGATATTAATTT 512
QY 161 ValGlyPheIlyPheThrValIlyIlyAspIlyAsnGluIlySerSerAspThrValGlnLeu 180
Db GTTGGTTTAATTCAGGATGAAGAAGATAAATAAGAAAGTCAGATCTGTCCAACTA 572
QY 181 LeuThrValIlyIlySerAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis 200
Db CTCACACTGCAAGAGATAGAGCGAAGGGCGCAGAACCTCTGTCCGAGCTGGAGACCA 632
QY 201 GlnGluProSerValGluThrAlaValGlyIlySerAlaSerIlySerGluLeuIly 220
Db CAGGAGCCGAGTGTGAGACTGCGTGGAGAGATCAGATCCAAAGAAAGTAGCTGAAG 692
QY 221 GlnSerThrGluIlySerGlnGluIlyThrLeuIlySerGlnGluIlySerSerThrGluIlePro 240
Db CAAATCCACAGAGAGCAAGAACACCTCGAAGCAAGAACAGACGACAGAAAATCCCC 752
QY 241 LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaIlySerAspGluGlyGluGluIly 260
Db CTTCAGGCGCAATCTGATCAAGCGCTGAGGAAGAACCAAGATCAAGAGAGAAAAA 812
QY 261 GlnGluIlyGluProThrIlySerSerProGluIlySerSerProValAsnSerGluThr 280
Db CAGAGAAAAGAGCCCAAGATCCCAAGATCCCGAGAGGCCAGCTCAACATGTAGACA 872
QY 281 ThrSerSerPheIlySerPhePheThrHisGlyTrpAlaGlyTrpArgIlyIlyThrSer 300
Db ACATCTTCTTCAAGAAGTCTTCACTCACTGAGTGGCGGCTGGGCAAGAGACAG 932
QY 301 PheIlyIlySerIlyGluAspAspLeuGluThrAlaGluIlySerArgIlySerGluGlnAla 320
Db TTCAGAAATCAAAAAGAGATGATCTGGAAACTGCCAGAGAAAGAAAGGAGCGCA 992
QY 321 GlnIlyValAspGluGluGluIlySerGluIlySerThrGluProAlaSerGluGluGlnIlyPro 340

Db 993 GAAAAAGTAGACGAGGAAGAAAAGAAAAGACAGAGCCAGCTCGGAGAGACAGAGCCG 1052
QY 341 AlaGluAspThrAspGlnAlaArgLeuSerAlaAspIlyGluIlyValGluLeuProLeu 360
Db GCAGAAACACAGACCAAGCCAGGTTGTCAGACACTACAGAAAGGTGAGACTGCTTTG 1112
QY 361 GluAspGluValAlaAspLeuGluAlaSerSerGluGluIlyIlySerValProLeuAlaThr 380
Db GAAGACAGGTTGGTACCTGAGGCATGTCAGAGAGAAAGTGTGCTCTTGGCAACG 1172
QY 381 GluValPheAspGluIlySerMetGluAlaHisGlnGluValValAlaGluValHisValSer 400
Db GAAGTGTGATGAGAGATGAGAAAGCCACCAAGAAGTGTGTCAGAGGTCCAGTGAGC 1222
QY 401 ThrValGluIlySerThrGluGluGlnGlyGlyGlyGluAlaGluGlyValVal 420
Db ACCGTGAGAAAGACAGAGAGAGACAGAGAGAGAGAGAGAGGTGAAGGGGCGTGTG 1292
QY 421 ValGluGlyThrGluIlySerLeuProGluIlyIlySerValAlaGluProGlnGluValPro 440
Db GTAGAAAGAACAGAGAAATCCTTGCCCTGAGAACTGGCTGAGCCCGAGAGGTCCC 1352
QY 441 GlnGluAlaGluProAlaGluGluIlyMetIlySerSerArgGluMetCysValSerGlyGly 460
Db CAGGAAGCTGAGCTGTGAGAGAGCTGATGAAGAGCAGAGAGATGTGTCTTGAGAGA 1412
QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluIlySerThrLeuProIlyHisProGlu 480
Db GACCACTCACTCACTCAGACCTTAAGTCTGAGAGAAAGAGCGCTGCCAAACACCAADA 1472
QY 481 GlyIleValSerGluValGluMetLeuSerSerGlnGluAlaGlyIleValGlnIlySer 500
Db GCATTTGTCAGTGAAGTGAAGTGTCTGCTTCAGGAAGAAATCAAGCTACAGGAAGT 1532
QY 501 ProLeuIlyIlyLeuPheSerSerSerGlyLeuIlyIlySerGlyIlyIlySerGlnIly 520
Db CCTTGAGAAAGAACTCTTCAGTATAGCTCAGGCTTTAAAGAGCTGTGGAGAAAGCAGAAG 1592
QY 521 GlyIlySerArgIlyGlyGlyIlyAspGluGluProGluIlyIlyThrGlnHisIleHisThrGlu 540
Db GGGAAACAGAGAGGTGGGGAGACAGAAAGCCTGAGAAATCAACATTCACACCCAA 1652
QY 541 SerProGluSerAlaAspGluGlnIlySerIlyIlySerSerAlaSerSerProGluIlyPro 560
Db TCCCAAGAGAGTGTGATGAGCAGAAAGGAGAGAGACTGTGGTGTCTCCCGAGGAGCTT 1712
QY 561 GlnGluThrThrCysLeuGluIlySerGlyProLeuGluAlaProGlnAspGlyGluAlaGlu 580
Db GAGAGACCACTGTGTGAGAAAAGGCCGCTGAAAGCAACCCAGATGGGGAAGCTGAG 1772
QY 581 GlnGlyThrThrSerAspGlyGluIlyIlySerArgGluIlyIleThrProTrpAlaSerPhe 600
Db GAAAGAACTACTTCCGATGGAGAGAAAGAGAGAGAGATCACTCCCTGGGCACTTTC 1832
QY 601 IlyIlySerMetValThrProIlyIlySerArgValArgAspProSerGluSerAspIlySerGlu 620
Db AAAAAGATGTGACACCCAGAAAGCGGTCCGAAACCTTCTGAGAGTGAACAAGAGAA 1892
QY 621 GluLeuGluIlyValIlySerSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
Db GAGCTGGAGAAAGTCAAGAGGCCCACTTGTCTCCTCACTGATAGCAGCTGCAGAAAATG 1952
QY 641 GluAspGluValIlySerThrValGlyGluGluGluIlySerProGluGluProIlyArgArgVal 660
Db CAAGATGAAGTCAAAACTGTTGGTGAAGAACAAAAGCCAGAGAAACCAAGCGTGGGTG 2012
QY 661 AspThrSerValSerThrGluAlaLeuIleCysValGlySerSerIlyIlyArgArgAlaArg 680
Db GATACTTCAGTGTCTTGGAGACACTGATTTGTGTGATATCTCCAAAGAGAGACCAAGG 2072
QY 681 IlyAlaSerSerSerAspAspGluGlyGlyProArgThrLeuGlyIlyIlyAspSerHisArg 700
Db AAGGATCTCTTCAAGATGATGAAGAGGGCCAAAGACACTGGAGGGGACAGTACAGA 2132

QY 701 AlaGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSerThrGln 720
DB 2133 GCAGAGAGCCAGCAAGAACAAAGAGCCGGAACAGACAGCTGTCTCCAGCAGCCAG 2192
QY 721 GluGlnAspGlnAlaGlnGlySerSerProGluProAlaGlySerProSerGluGly 740
DB 2193 GAGCAGGACCAAGCCAGAGAGTTCTCCAGCAGCCAGGAGCCCTCCGAAAGG 2252
QY 741 GluGlyValSerThrTrpGluSerPheLysArgLysValThrProArgLysLysSerLys 760
DB 2253 GAAGGTGTCTCCACTTGGGAGTCATTTAAAGATTAGTCACTCCAAGAAAAAATCCAAG 2312
QY 761 SerLysLeuGluGlnLysAlaGluAspSerSerValGluGlnLeuSerThrGluLeuGlu 780
DB 2313 TCAAAACTGGAAGAGAAAGCCGAGACTCTAGTGTAGAGCAGTTGTCCACTGAGATCGAA 2372
QY 781 ProSerArgGluGluSerTrpValSerIleLysLysPheIleProGlyArgArgLysLys 800
DB 2373 CCAGGTAGAGAAATCTTGGGTTTCCATTAGAAATTCATCCCCGACCGCGGAAGAAA 2432
QY 801 ArgAlaAspGlyLysGlnGluAlaThrValGluAspSerGlyProValGluIleAsn 820
DB 2433 AGGCGAGCCGGAAGCAAGAACCAAGCCACTGTGGAAGACTCAGGCCAGTGGAGATAAT 2492
QY 821 GluAspAspProAsnValProAlaValValProLeuSerGluTrpAsnAlaValGluArg 840
DB 2493 GAGGACAGCCCTAATGTCCAGCCGCTCGTGCTCTGAGTATATGCAATGTCAGTGGAGG 2552
QY 841 GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrVal 860
DB 2553 GAGAAGTGAAGCCCAAGGGGAATACGGAGCTGCCAGCTGTGGGGGCTGTGTACGTG 2612
QY 861 SerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValIleAspGlyThr 880
DB 2613 TCCGAGGAGCTCAGTAGACTCTGGTCCACACTGTGAGTGTGCGAGTCATTGTGGACC 2672
QY 881 ArgAlaValThrSerValGluGluArgSerProSerTrpLysSerAlaSerValThrGlu 900
DB 2673 AGGCGAGTCAACAGTGTGGAAGAGCGGTCTCTTCTGCGGATATCCGCTCCGTAACAGAA 2732
QY 901 ProLeuGluHisThrAlaGlyGluAlaMetProProValGluGluValThrGluLysAsp 920
DB 2733 CCTCTGAACACACAGCGGAGAGCCATGCCACCTGTTGAAGAGGTCACTGAAAAAGAC 2792
QY 921 IleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
DB 2793 ATCATTCAGAGAAACTCTGTGCTCACCAGACGTTACAGAGGGTAAAGATGCCAT 2852
QY 941 AspAspMetValThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThr 960
DB 2853 GAGCAGATGGTCAACAGTGAAGTGAATTTCACTCAGAAAGCTGTGACAGCCACAGAGACC 2912
QY 961 SerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAsp 980
DB 2913 TCAGAGGCTCTCCGTACTGAAAGATTACCGAAGCATTCGGGGCCGGAAGACCCACAGAC 2972
QY 981 MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrPro 1000
DB 2973 ATGGTGTCCGAGTTTCCAGTGTACTGTCTCCAGACACCCACAGAGGAAGCCACCCCA 3032
QY 1001 ValGlnGluValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAla 1020
DB 3033 GTTCAGGAGGTAGAGAGTGGTGTCTAGATACAGAAAGAGAGGCGCCAGAGCCAGGCC 3092
QY 1021 IleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThr 1040
DB 3093 ATCCTCCAAGCCGTGAGCAACAAGGTGAAAGAGGAGTCCAGAGTGCCTGCAACCCAGACT 3152
QY 1041 ValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGluAspSerGlu 1060
DB 3153 GTCCAGAGAAACGGGTCAAAAGCACTGGAGAGGTGAGGAGGTAGAGGAGTCCGGA 3212

QY 1061 ValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGlnGluAlaGly 1080
DB 3213 GTGCTGGCTTCGAGAGAAAGAGAGGACGTATATGCCGAAAGAGCCCGTCAGGAAGCTGA 3272
QY 1081 AlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluVal 1100
DB 3273 GCTGAGCATCTTGCACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCTTGAAGTT 3332
QY 1101 ProGluValThrAlaAspValAspHisValAlaThrCysGlnValIleLysLeuGlnGln 1120
DB 3333 CCTGAAGTCACGCGCAGATGTAGACCATGTGCCACGTCGCGAGGTATCAAGCTCCAGCAG 3392
QY 1121 LeuMetGluGlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsn 1140
DB 3393 CTGATCGAACAGGCGCTGGCCCTGAGTTCATCCGAAACCTTGACACACAGTGAACAAAT 3452
QY 1141 GlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIle 1160
DB 3453 GGNAGCACTCCCTTAGCAGATTGACACACTGCAGATGGGACACAGCAAGATGAACCAAT 3512
QY 1161 AspSerGlnAspSerLysAlaThrAlaValArgGlnSerGlnValThrGluGluGlu 1180
DB 3513 GACAGCCAGGACAGTAAAGCCACTGCAGCTGTCCAGGCTGCAGGCTCAGCAAGAAAGAG 3572
QY 1181 AlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGlu 1200
DB 3573 GCGGCTACTGCTCAGAAAGAGGAGCCTTCGACACTACCTAATAATTTCCAGGCCAGGAA 3632
QY 1201 GluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAla 1220
DB 3633 GAACATGGGGAAGAACCCAGAAAGAGATGTTCTTGAACCTTACACAGCAAGAGCTTACTGCT 3692
QY 1221 AlaAlaValProValLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeu 1240
DB 3693 GCAGCCGTGCCGCTTCTGCAAAAGACTGAGGTGGGTCAAGAGGGTCAAGTGTGCTGCTG 3752
QY 1241 AspGlyGluLysValLysGluGlnGluValPheValHisSerGlyProAsnSerGln 1260
DB 3753 GATGGAGAAAAAGTCAAAAGAACAGGAGGTGTTTTGTACACTCTGACCCCAACAGTCAA 3812
QY 1261 LysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLys 1280
DB 3813 AAGCTCTGTGTGTGACATATGACATGAGTGAAGTGTGGGAGTGGCCGGGTGTCAAGAAAAG 3872
QY 1281 GluSerThrGluValGlnSerLeuSerLeuGluGlyGlyMetGluThrAspValGlu 1300
DB 3873 GAGAGTACTGAAGTCAGAGCTCTTACCTCGAGGAGGAGAGATGAAACTGACGTTGAA 3932
QY 1301 LysGluLysArgGluThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAla 1320
DB 3933 AAGGAGAAAAGGAGACAAAAGCCAGAGCAAGTGAAGTGAAGAGGTGAGCAGGAAACAGCC 3992
QY 1321 AlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSer 1340
DB 3993 GCTCCTGAGATGAAGAAACCTACGGGAAGCCAGTCTGACACTTGACATGCCAGCTCA 4052
QY 1341 GluArgGlyLysAlaLeuGlySerLeuGlyLysGlySerProSerLeuProAspGlnAspLys 1360
DB 4053 GAGAGGGGAGAGCACTGGGAGCCCTTGGAGGAGCCCTTCTCTCCAGCAACAGACAAA 4112
QY 1361 AlaGlyCysIleGluValGlnValGlnSerLeuAspThrValThrGlnThrAlaGlu 1380
DB 4113 GCAGGTTCATAGAGGTTCAAGTTCAAGCCTGGAACACACAGTCACTCAACAGCAGAA 4172
QY 1381 AlaValGluLysValIleGluThrValIleSerGluThrGlyGluSerProGluCys 1400
DB 4173 GCTGTGAAAAGGTCTATAGAAACGGTTGTGATTTTCAGACAGAGGTGAAAGTCCAGAGTGT 4232
QY 1401 ValGlyAlaHisLeuLeuProAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeu 1420
DB 4233 GTAGGTGCACACTTATACAGCTGAGAGTCTCTCTGCAACGGGTGGCCACTGACTCTT 4292
QY 1421 GlnHisAlaGluAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProfile 1440

Db 4293 CAGCATGACAGAGACCGGTACCCCTGGGACCTGACTCTCAAGCAGAAATCCATCCCAATC 4352
Qy 1441 lIeValThrProAlaProGluSerThrLeuHisProAspLeuGlnGlyLuiIeSerAla 1460
Db 4353 AATAGTAACTCCCTGCTCTTAAAGACACCTTACATCTTGACCTTCAAGAGAAATTAAGCGCA 4412
Qy 1461 SerGlnArgGluArgSerGluGluIuAspLysProAspAlaGlyProAspAlaAspGly 1480
Db 4413 TCCCGAGAGAGACGATCAAGAGAGAGAGCAAGCCAGATGCTGGTCTTATGCTGACGGC 4472
Qy 1481 LysGluSerThrAlaIleGluLysValLeuLysAlaGluProGluLuiLeuGluLeuGlu 1500
Db 4473 AAGGAGAGTACAGCAATTCGAAAAGTCTTCAAGGCTGAACCTGAGATCCTGGAACTTGAG 4532
Qy 1501 SerLysSerAsnLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArg 1520
Db 4533 AATTAGAGCAACAAGTTGTGTGGAACGTTCATTCAAGACGCCGTTGACCAAGTTCCGACGT 4592
Qy 1521 ThrGluThrAlaProGluThrHisAlaIlyrAspSerGlnThrGlnValProAlaCybArg 1540
Db 4593 ACAGAAACAGCCCCCGAAATCATGCTTATGATTCACAGACCCAGGTTCTCGATGCGACG 4652
Qy 1541 LeuAspSerArgGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHis 1560
Db 4653 CTTGACAGACAGGAGCCCAACAGATCTGACAAAATGAAAGATGCCAAGATGAACAC 4712
Qy 1561 ProValProGlnProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGln 1580
Db 4713 CCAGTCCCGCCAGCCGCAAGAGAGACTTGCAAGTCTTGACCGTTCTGAGGAGATGGGCTCAG 4772
Qy 1581 ProArgLysCysLeuProArgLeuGlnLeuLysAlaProValSerLys 1596
Db 4773 CCTCGGAAATGCTTGCGCGCTTGCAATTGAAGACGCCGATGTCAAG 4820

RESULT 2
US-09-902-432-1
; Sequence 1, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Susan G. Jakem
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156, 0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5134
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-902-432-1

Alignment Scores:
Pred. No.: 0 Length: 5134
Score: 6005.00 Matches: 1232
Percent Similarity: 97.33% Conservative: 7
Best Local Similarity: 96.78% Mismatches: 27
Query Match: 74.38% Indels: 14
DB: 9 Gaps: 4

US-09-902-432-4 (1-1596) x US-09-902-432-1 (1-5134)
Qy 329 GtLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArg 348
Db 2 GAAAGACAGAGCGCTCGAGAGAGAGAGAGAGCCGCGCAAGACACAGACCGCCAGG 61

Qy 349 LeuSerAlaAspTrpGluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGlu 368
Db 62 TTGTACAGAGACTTACAGAGAGGTGAGCTGCTTGGAAAGACAGGTGTGACCTTGAG 121
Qy 369 AlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLysMetGlu 388
Db 122 GCATGCTCAGAGAGAGAGATGTGCTCTTGGCACCGAAGTGTGATGAGAAAGTGA 181
Qy 389 AlaHisGlnGluValAlaGluValHisValSerThrValGluLysThrGluGluGlu 408
Db 182 GCCCACCAAGAGTGTGTCAGAGAGGTCCAGGTGACACCTGAGAAAGACAGAGAGAG 241
Qy 409 GlnGlyGlyGlyGlyGluAlaGluGlyGlyValValAlaGluGlyThrGlyGluSerLeu 428
Db 242 CAGGAGAGAGAGAGAGAGCTGAAGGGGGCTGTGTGTAGAGAAACAGAGAAATCTTGG 301
Qy 429 ProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGlu 448
Db 302 CCCCTTGAAACTGGCTGAGCCCCAGAGAGTCCCCAGAAAGCTGAGCTGTGAGAG 361
Qy 449 LeuMetLysSerArgGluMetCysValSerGlyIAspHisThrGlnLeuThrAspLeu 468
Db 362 CTGATGAAGAGAGAGAGATGTGTCTCTGAGAGAGACCACTCACTGACAGACCTA 421
Qy 469 SerProGluGluLysThrLeuProLysHisProGluGlyIleValSerGluValGluMet 488
Db 422 AGTCCCTGAAGAGAAACCTGCCCCAACCAGAAAGCATTTGTCAGAGGTGAGATG 481
Qy 489 LeuSerSerGlnGluArgIleLysValGlnGlySerProLysLysLeuPheSerSer 508
Db 482 CTGTCTCTCAGAAAGAAATCAAGGTACAGGGAAGTCCCTTGAAAGAACTTTCAGTAC 541
Qy 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLys 528
Db 542 TCAGGCTTAAAGAACTGTCTGCGAAGACAGAAAGGGAACAGAGAGGTGGGAGAC 601
Qy 529 GluGluProGluGlyLysThrGlnHisIleHisThrGluSerProGluSerAlaAspGluGln 548
Db 602 GAAGAGCTCGAAGATACCAACATTCACACCGAATCCCAAGAGAGTGTGATGACAG 661
Qy 549 LysGluLysSerSerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLys 568
Db 662 AAGGAGAGAGCTTGGCTGTCTCCCGAGAGCTTGAGAGACCACTGTGAGAA 721
Qy 569 GlyProLeuGluAlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGlu 588
Db 722 GGGCCGCTGAAAGACCCAG-GATGGGAGCTGAGAAAGAACTACTTGGT -GGAAG 778
Qy 589 LysLysArgGluGlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLys 608
Db 779 AAGAAGAGAGAGG--ATCATCTCCCTGGGCATCTTCAAAAAGATGTGACACCCAAAGAA 836
Qy 609 ArgValArgArgProSerGluSerAspLysGluGluGluLysValLysSerAla 628
Db 837 CGGTC-CGAAAGCTTCTGAGAGTACAAAGAGAGAGCTGGAAGAGTCAAGAGCC 895
Qy 629 ThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGly 648
Db 896 ACCTGTCTCTCACTGATAGACACAGTGTCAAGAAATGCAAGATGAAGTCAAAATCTTGGT 955
Qy 649 GluGluGlnLysProGluGluProLysArgValAspThrSerValSerTrpGluAla 668
Db 956 GAGGAACAAAGCCAGAGAAACCAAGCGTGGGTGATCTTCACTGCTTGGAGAAC 1015
Qy 669 LeuLysCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGlu 688
Db 1016 CTGATTTGTGTGATCATCCAAAGAGAGAGAGAGAGGCACTCTTCAAGATTTAAGA 1075
Qy 689 GlyGlyProArgThrLeuGlyLysGlyAspSerHisArgAlaGluGluAlaSerLysAspLys 708
Db 1076 ---GGGCAAGGACACTGGAGGGGGGACAGTCAACAGACAGAGAGCCAGCAAGAAAGACA 1132

QY 708 sGluAlaGlyThrAspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySe 728
Db 1133 AGAAGCCGA-ACAGACGCTGTTCTGCGACACCCAGAGCAGGACCAAGCGCAAGGAAG 1191
QY 728 rSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSe 748
Db 1192 TTCTCACCAGGAGCAGCGGAGCCCTTCGAAAGGGAAGGTGCTCCACTTGGGAGTC 1251
QY 748 rPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGluLysAla-- 767
Db 1252 ATTTAAAGATTAGTCACTCCAGAAAATAATCCAAGTCAAAACTGGAAGAGAAGA-AG 1310
QY 768 ----GluAspSerSerVal---GluGlnLeuSerThrGluLleGluProSerArgGluG1 785
Db 1311 CCGAAGGACTCTAGTTGTAGGACGAGCTTGTCCACTGAGATCGAACCGGTGAGAAGA 1370
QY 785 uSerTrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLY 805
Db 1371 ATCTTGGGTTTCCATTAAAGAAATTCATCCCGGACGCGGAAAGAAAGGCGAGATGGAA 1430
QY 805 s-GlnGluGlnAlaThrValGluAspSerGlyProValGluLleAsnGluAspAspProA 825
Db 1431 GGCAAGAACCAAGCCACTGTGGAAGACTCAGGGCCAGTGGAGATAAATGAGACGAGCCTG 1490
QY 825 snValProAlaValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGluA 845
Db 1491 ATGTCCAGCGCTGTCCTGCTGAGTATGATCAGTGGAGGAGGAGAAATGATGAAG 1550
QY 845 laGlnGlyAsnThrGluLeuProGlnLeuGlyAlaValTyrValSerGluGluLeuS 865
Db 1551 CCCAGGGGAATGCGGAGCTGCCAGCTGCTGGGCTGTGTA--GTGTCGAGGAGCTCA 1607
QY 865 erLysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrS 885
Db 1608 GTAAGACTCTGGTCCACACTGTGAGTGTGCGAGTCAATTGATGGACCCAGGCGAGTCA 1667
QY 885 erValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisT 905
Db 1668 GTGTGGAAGAGCGGTCTCTTGTGGATATCCGCTTCGTAACAGAACCTCTTTGAAACACA 1727
QY 905 hrAlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAlaGluG 925
Db 1728 CAGCGGAGAAGCCATGCCACTGTTCAGAGGTCCTGAAAGAGACATCATTTGCAGAAG 1787
QY 925 luThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetValT 945
Db 1788 AAACCTCTGTGCTCACCCAGACGTTTACAGAGGTAAGATGCCATGACCAATGGTCA 1847
QY 945 hrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuA 965
Db 1848 CCAGTGAAGTGGATTTTCACTCAGAAGCTGTGACAGCCACAGAGACCTCAGAGGCTCTCC 1907
QY 965 rgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlav 985
Db 1908 GTACTGAAGAAGTTTACCGAAGCATCGGGCCGAGAGACCAACACATGGTGTCCGAG 1967
QY 985 alSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluValG 1005
Db 1968 TTTCCAGCTGACTGACTCTCCACACACACACAGAGGAAGCCACCCAGTTTCAGGAGTAG 2027
QY 1005 luSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAlav 1025
Db 2028 AGGTGTGTGTGTAGTACAGAGAAGAGGAGCGCCAGAGCGCCAGGCGCATCTCCAAAGCCG 2087
QY 1025 alAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThrG 1045
Db 2088 TTGCAGACAGGTGAAGAGAGGTCCCGAGGTGCTGCAACCCAGACTGTGCAGAGAACGG 2147
QY 1045 lySerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAlaSerG 1065
Db 2148 GGTCAAAAGCACTGGAGAAGGTTGAGGAGGTAGAGGAGGACTCCGAAGTGTGCTTCGG 2207
QY 1065 luLysGluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuA 1085

Db 2208 AGAAAGAGAAGGACGTTATGCCGAAAGGACCCGTGCAGGAAGCTGGAGCTGACATCTTG 2267
QY 1085 laGlnGlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrA 1105
Db 2268 CACAGGGCTCTGAGACTGGACAGGCTACTCCAGAGAGCCCTTGAAGTTCCTGAAGTCACAG 2327
QY 1105 laAspValAspHisValAlaAlaThrCysGlnValIleLysLeuGlnLeuMetGluGlnA 1125
Db 2328 CAGATGTAGACCATGTGCGCACGTCGAGGTATCAAGCTCCAGCAGCTGATGAAACAGG 2387
QY 1125 laValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProL 1145
Db 2388 CCGTGCCCTTGAGTCATCCGAAACCTTGACAGACAGTGAACAAATGGAAGCACTCCCT 2447
QY 1145 euAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAsps 1165
Db 2448 TAGCAGATTTCAGACACTGCAGATGGGACACAGCAAGATGAACCATTCAGCAGCAGACA 2507
QY 1165 erLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaG 1185
Db 2508 GTAAGCCCACTGCAGCTGTGAGCAGCTCAAGGTCAAGAGGAGGCGGCTACTGCTC 2567
QY 1185 lnLysGluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHisGlyGluG 1205
Db 2568 AGAAAGAGGAGCCCTTCGACACTACCTAAATTAATGTTCCAGCCAGGAAGAACATGGGGAAG 2627
QY 1205 luProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaValProV 1225
Db 2628 AACCCAGAGAGAGATGTTCTTGAACCTACACAGCAAGAGCTTGTCTGCAGCCGCTGCCG 2687
QY 1225 alLeuAlaLysThrGluValGlyGlnGluGlyGluValAspTrpLeuAspGlyGluLysV 1245
Db 2688 TCTGCGAAAGAGACTGAGGTGGGTCAAGAGGTGAGGTGACTGTTGGATGGAGAAAAG 2747
QY 1245 alLysGluGlnGluValPheValHisSerClyProAsnSerGlnLysAlaAlaAspV 1265
Db 2748 TCAAGAAAGAACAGAGGAGGTGTTGTACACTCTGGACCCCAACAGTCAAAAGGCTGCTGATG 2807
QY 1265 alThrTyrAspSerGluValMetGlyValAlaGlyCysGlnGluLysGluSerThrGluV 1285
Db 2808 TGACATATCACAGTGAAGTGATGGGAGTGGCCGGTGTCCAGGAAAGAGAGAGTACTGNAAG 2867
QY 1285 alGlnSerLeuSerLeuGluGluGlyGluMetGluThrAspValGluLysGluLysArgG 1305
Db 2868 TGACAGAGTCTTAGCTCTGGAGGAGGAGAGATGGAATCTGCTGTAAGAGGAGAGAAAGGG 2927
QY 1305 luThrLysProGluGlnValSerGluGluGlyGluGlnGluThrAlaAlaProGluHisG 1325
Db 2928 AGACAAAGCCAGAGCAAGTGAGTGAAGAGGTGAGCAGGAACACAGCCGCTCTCTGAGCATG 2987
QY 1325 luGlyThrTyrGlyLysProValLeuThrLeuAspMetProSerSerGluAtgGlyLysA 1345
Db 2988 AAAGGAACCTACGGGAAGCCAGTCTCTGACACTTGACATGCCAGCTCAGAGAGGGGAGAGG 3047
QY 1345 laLeuGlySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleG 1365
Db 3048 CACTGGGAAGCCTTGGAGGAAGCCCTTCTCTCCAGACCAACAGCAAGCAGAGTTCATAG 3107
QY 1365 luValGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysV 1385
Db 3108 AGGTTCAGATGTTCAAGCCTGGACACAACTGCTCAAAACAGCAGGAAGCTGTGGAAAGG 3167
QY 1385 alIleGluThrValValIleSerGluThrGlyLysSerProGluCysValGlyAlaHisL 1405
Db 3168 TCNTAGAAACGGTGTGATTTTCAGACAGAGGTGAAGTCCAGAGTGTGTAGGTGCACACT 3227
QY 1405 euLeuProAlaGluLysSerSerAlaThrGlyHisTrpThrLeuGlnHisAlaGluA 1425
Db 3228 TATTACCACTGAGAAGTCTCTCTGCAACGGGTGGCCACTTGGACTCTTCAGCATGCAGAGG 3287
QY 1425 spThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrProA 1445

Db 3288 ACAAGTACCCCTGGGCGCTGAGTCTCAGGACGAATCCATCCCAATCATAGTAACTCTCG 3347
Qy 1445 IAPROGLUSERThRleuHISProAspLeuGlnGlyGluIleSerAlaSerGlnAlaGlnAla 1465
Db 3348 CTCCTGAAAGACACCTCACTCTGACCTCAAGAGAAATAAGCCATCCAGAGAGAGC 3407
Qy 1465 rgsrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1485
Db 3408 GATCAG 3467
Qy 1485 Ia11GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1505
Db 3468 CAATGACAAAGTCTCTCAAGGCTGAACTGAGATCTCTGAACTTGAAGATTAAGAGACA 3527
Qy 1505 ys11eValleuAnVal11egInThAlaValAspGlnPheAlaGlnGlnGlnGlnGln 1525
Db 3528 ACATTGTGCTGAACGTCATTCAGACAGCCGTTGACCAAGTCCGACGCAAGAAACGCC 3587
Qy 1525 rGlnGlnThHISAlaThraAspSerGlnThrGlnValProAlaCysArgLeuAspSerArg 1545
Db 3588 CCGAAACTCATGCTTAATGATTCACAGACCCAGGTTCTCTCAATGCGCTTGAACAGAGAG 3647
Qy 1545 IAPROAsnArgCysTrpThrLysMetLysAspAlaLysMetLysHisProValProGlnP 1565
Db 3648 AGCCCAACAGATGCTCTGACAAATAATGAAGTTGCCAAGATGAACACCCAGTCCGACAG 3707
Qy 1565 rGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1585
Db 3708 CCAGAGAGAGACTTGCAAGTCTGACCGCTTCTGAGAGGATGCG--TCAGCTCGGAATGCT 3765
Qy 1585 euProArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1596
Db 3766 TGCCGCGCTTGACATTGAAGCGCCGCTGTCAAG 3800

RESULT 3

US-09-738-877-2
; Sequence 2, Application US/09738877
; Patent No. US2002015970A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan
; APPLICANT: Weiss, Stephen J.
; APPLICANT: Glynn, Richard
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF ANGIOGENESIS, COMPOSITIONS, AND MEI
; FILE REFERENCE: A-69806/DJB/JJD
; CURRENT APPLICATION NUMBER: US/09/738,877
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT/US 00/22061
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patencin version 3.1
; SEQ ID NO 2
; LENGTH: 5346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-877-2

Alignment Scores:

Pred. No.: 1,22e-292 Length: 5346
Score: 4300.50 Matches: 967
Percent Similarity: 68.81% Conservative: 189
Best Local Similarity: 57.56% Mismatches: 408
Query Match: 53.27% Indels: 116
DB: 10 Gaps: 36

US-09-902-432-4 (1-1596) x US-09-738-877-2 (1-5346)

Qy 1 MetGlyAlaGlySerThrGlnGlnArgSerProGlnGln--ProAlaGlySerAap 19

Db 1 ATGGGCGCGGGAGAGCTCCACCGAGCAGCGCAGCCGAGAGAGCCCGGAGGAGAGCTCC 60
Qy 20 ThrProSerGlnLeuValLeuSerGlyHisGlyProAlaAlaGlnAlaSer--GlyAla 38
Db 61 ACGCGGCTGAGCCCGAGCCAGCCAGCCGCGCGCGCTGCGCGGAGAGAGAGAGAGAGAG 120
Qy 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
Db 121 ACGCGGAGCCCGCGATGCTGCTCTCGGACCCGCGACCAAGCTCTTACAGAAAGATGCT 180
Qy 57 GlnLeuSerSerValAsnGlyValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 76
Db 181 CAGCTGTCCACCATCAATAGGGGTAGCTAGCAATAGAGCTCAGCCCTCAGAGAGGTAC 240
Qy 77 GlnGlnGlnGln-----GlnGln 82
Db 241 CTAAATGGCCAGAAAGAGAGCCCTGAACGCTCAAGAGAGCCCTAAACAGCCAGAGAGAA 300
Qy 83 GlnValValAspGlnAspValGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 102
Db 301 GAAGTCATGTCACGAGAGGTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy 103 ValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 122
Db 361 GATAAAGAGATGCTACTAAGTCAAGCTGTTTCAAGACATCAAGATGATGGCAGAG 420
Qy 123 GlnThrSerGlnLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 142
Db 421 GAGAACCGAAAT--ATCGAACAGATCTCTTCAAGAAACAAATTTGAGAGAGCTACA 477
Qy 143 GlnProAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 162
Db 478 CAACCCACTGAGTCCAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 537
Qy 163 PheLysPheThrValLysLysAspLysAsnGlnLysSerSerAspThrValGlnLeuLeuThr 182
Db 538 TTTAAATTCACGTGTAAGAAAGATGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
Qy 183 ValLysLysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 202
Db 598 GTGAAGAAAGATGAAG 648
Qy 203 ProSerValGlnThrAlaValGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 222
Db 649 CCCAGCCTT-----GGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
Qy 223 ThrGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 242
Db 703 ACAGAGAAACCGAAG 762
Qy 243 AlaGlnSerAspGlnAlaAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 262
Db 763 GCCGATCTGGCCAG 819
Qy 263 LysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 282
Db 820 AAAAGAACTAGCAAGTCTGCAAGAACTCCGACTACCTCCGAGAGAGAGAGAGAGAG 879
Qy 283 SerPheLysPhePheThrHisGlyTrpAlaGlyTrpAlaGlyTrpAlaGlyTrpAlaGly 302
Db 880 ACCTTCAAAATTTCTTCACTCAAGGTGGCGGCTGCGGAGAGAGAGAGAGAGAGAG 939
Qy 303 LysSerLysGlnAspAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 322
Db 940 AAGCCGAAG 999
Qy 323 ValAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 337
Db 1000 GTAGACACAGAAAG 1059
Qy 338 -----GlnGlnProAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 353

Db 1060 CAAGCCACCACAGGAGCGGCAGAAAGTGCCTCCACAGAGCCCGGTTATCAGCTGAATAT 1119
Qy 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
Db 1120 GAGAAAGTTGAGCTGCCCTCAGAGGAGCAAGTCAGTGGCTCGCAGGGACCTTCTGAAGAG 1179
Qy 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
Db 1180 AAACCTGCTCCGTGGCGCAGAGAAGTGTGATGAGAAAAATAGAGTCCACCAAGAGAG 1239
Qy 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGluGlyGly 412
Db 1240 GTTGTGGCCGAGTCCACGTCAGCACCGTCGAGAGAGAACCGAGAGCAG----- 1290
Qy 413 GlyGluAlaGluGlyGlyValValGluValGluGlyThrGlyGluSerLeuProGluLys 432
Db 1291 -----AAAACGGAGGTGGAAGAAACAGCAGGGTCTGTGCCAGCTGAAGAA 1335
Qy 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLysSer 452
Db 1336 TTGGTTGGAATGGATGCAGAACCTCAGGAAGCCCAACCTCCCAAGGAGCTGGTCAAGCTC 1395
Qy 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
Db 1396 AAAGAAACGTGTGTTCCGAGAGGACCCTACAGAGGAGCTGACCTCAGTCTCTGATGAG 1455
Qy 473 LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGln 492
Db 1456 AAGTGTGTGCCAAACCCCGAAGGGGTTGTGAGTGAGGTGGAATGCTGTCATCACAG 1515
Qy 493 GluArgIleLysValGlnGlySerProLeuLysLysLeuPheSerSerGlyLeuLys 512
Db 1516 GAGAGATGAAGTGCAGGGAGTCCACTAAAGAAGCTTTTACCAGCACTGGCGTTAAAA 1575
Qy 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyValAspGluGluProGly 532
Db 1576 AAGCTTTCTGGAAGAAACAGAAAGGGAAGA---GGAGGAGAGCAGAGGAATCAGGG 1632
Qy 533 GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGlnLysGlyGluSer 552
Db 1633 GAGCACACTCAGGTTCACGCGGATTCTCCGACAGCCAGGAGGACCAAGAGGCGAGAGC 1692
Qy 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
Db 1693 TCTGCTCATCCCTGAGGAGCCCGAGGAGATCACGTGCTCTGGAAGAGGCGCTTAGCCGAG 1752
Qy 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
Db 1753 GTGCAGCAGGATGGGGAAGCTGAAGAGGAGCTACTTCCGATGGAGAGAAAGAGAGAA 1812
Qy 593 GlyIleThrProTrpAlaSerPheLysMetValThrProLysLysArgValArgArg 612
Db 1813 GGTGTCACTCCCTGGGCATCATTCAAAAAGATGGTGACGCCCAAGAGCGTGTAGACGG 1872
Qy 613 ProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
Db 1873 CCTTCGAAAGTGATAAAGATGAGCTGGACAAGGTCAAGGCTCAAGAGCGCTACTCTGTCTCC 1932
Qy 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluGlnLys 652
Db 1933 ACCGAGAGCAGCCTCTGAATGCAAGAGAAATGAAGGGAGCGGTGGAAGAGCCCAAG 1992
Qy 653 ProGluGluProLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
Db 1993 CCGAAGAACCAAGCCAGGTGGATACCTCAGTATCTTGGGAGCTTTAATTTGTGTG 2052
Qy 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGlyProArg 692
Db 2053 GGATCATCCAAGAAAAAGCAAGAGAGAGTCTCTTCTGATGAGGAAGGGGAGCCAAAA 2112
Qy 693 ThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
Db 2113 GCAATGGGAGGAGACCACCAAGAAAGCTGATGAGGCCCGGAAGAAAGCAAGAGACGGGACA 2172

Qy 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGlu 732
Db 2173 GACGGGATCCTTGCTGTTCCCAAGAAACATGATCCAGGGCAGGGAAGTCTCTCCCGGAG 2232
Qy 733 ProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGluSerPheLysArgLeu 752
Db 2233 CAAAGCTGGGAAGCCCTACCCGAAGGGGAGGCGTTTCCACCTGGGAGTCATTTAAAGGTTA 2292
Qy 753 ValThrProArgLysLysSerLysSerLysLeuGluGluLysAlaGluAsp----- 769
Db 2293 GTCACGCCCAAGAAAAAATCAAGTCCAGCTGGAAGAGAAAGCGAAGACTCCATAGCT 2352
Qy 770 ---SerSerValGluGlnLeuSerThrGluLeuGluProSerArgGluGluSerTrpVal 788
Db 2353 GGGTCTGGTGTAGAACATTTCCACTCCAGACACTGAAACCCGCTAAAGAAATCCTGGGTC 2412
Qy 789 SerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
Db 2413 TCAATCAAGAAGTTTATCTCTGACGAAGAAAGAACCCAGATGGGAAACAAAGAACAA 2472
Qy 809 AlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAla 828
Db 2473 GCCCTGTTGAAGACGACGAGGCCCAACAGGGGCCAACGAAGTGACTCTGTGTCCGGCC 2532
Qy 829 ValValProLeuSerGluTyrAsnAlaValGluArgGluLysMetGlu-----Ala 845
Db 2533 GTGGTCCCTCTGTCTCAGTATGATGCTGTAGAAAGGAGAGAAATGAGGCACAGCAAGCC 2592
Qy 846 GlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSer 865
Db 2593 CAAAAAGGCGCAGAGAGCCGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2652
Qy 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
Db 2653 GAGAGTCAGTTTATGATGATGAGCAGCAGCTGTCTGACGGAGGAGGAGGAGGAGGAGGAG 2712
Qy 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProLeuGluHisThr 905
Db 2713 ATTGAAGAAAGGTCTCTTCTTGATATCTGCTTCAAGTGCAGAACCTCTTGAACAAGTA 2772
Qy 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleAla---Glu 924
Db 2773 GAAGCTGGAAGCCGCACTGTTAACTGAGGAGGTATTGGAAGAGAGAAATTTGAGAGAA 2832
Qy 925 GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal 944
Db 2833 GAACCCCCACGGTTACTGAACTCTGCGAGAGAACAGAGAGCCCGGGCGGACACGGTC 2892
Qy 945 ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu 964
Db 2893 GTTAGTGAGCGGAATTGACCCCGGAGGCTGTGACAGCTGCAGAACTGCAGGGCCATTG 2952
Qy 965 ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla 984
Db 2953 GGTTCGAAAGAGAAACCGAAGCATCTGCTGTGAAGAGACACAGAAATTTGTTGTCAGCA 3012
Qy 985 ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal 1004
Db 3013 GTCTCCAGTTTAACCGACTCCCGCAGACACACAGAGAGGCGCACTCCGGTGCAGGAGGTG 3072
Qy 1005 GluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
Db 3073 GAAGTGGCGTACCTGACATAGAGAGCAAGAGAGCGGAGCTCAAGAGGTCTCCAGGCA 3132
Qy 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThr----- 1038
Db 3133 GTGGCAGAAAAAGTGAAGAGGAATCCAGCTGCTGCGACCGGTGGGCGCAGAAAGTGTG 3192
Qy 1039 ---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu 1057
Db 3193 CTTTACGCTGTGCAGAGACA-----GAGGCAGAAAGAACCCAGAGAG 3234

JOURNAL: Curr. Biol.
VOLUME: 7
ISSUE: 1
PAGES: 52-62
DATE: 1997-01-01
DATABASE ACCESSION NUMBER: U81607
DATABASE ENTRY DATE: 1997-06-26
US-09-738-877-1

Alignment Scores:

Pred. No.: 1,566-292 Length: 6608
Score: 4300.50 Matches: 967
Percent Similarity: 68.81% Conservative: 189
Best Local Similarity: 57.56% Mismatches: 408
Query Match: 53.27% Indels: 116
DB: 10 Gaps: 36

US-09-902-432-4 (1-1596) x US-09-738-877-1 (1-6608)

QY 1 MetGlyAlaGlySerThrGluArgSerProGluGln---ProAlaGlySerAsp 19
DB 192 ATGGGCGCCGGAGCTCCACGAGCAGCGCGCGGAGCAGCGCGCGGAGGAGCTCC 251
QY 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
DB 252 ACGGCGGCTGAGCCGAGCCGAGCGCGCGGCGCCCTCGCGCGGAGCGCGCGCAGACACC 311
QY 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
DB 312 ACCGCGGACCCCGCATCGCTCGCTCGGACCCCGCCACCAAGCTCTTACAGAGAAATGGT 371
QY 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsn 76
DB 372 CAGCTGTCCACCATCAATGCGTAGCTGAGCAAGATGAGCTCAGCCTCCAGGAGGGTGAC 431
QY 77 GlnGluGlyGln-----GluGlu 82
DB 432 CTAATGCGCAGAAAGAGAGCCCTGAACGGTCAAGAGCCCTTAACACGCCAGGAGGAAGAA 491
QY 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
DB 492 GAAGTCATTGTCAGGAGGTGGACAGAGAGACTCTGAAGATGTGACGGAAGAGACTCC 551
QY 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspThrLysAspGlyGlnGlu 122
DB 552 GATAAAGAGATGGCTACTAAGTCAGCGGTGTTTCCAGCATCATCAGATGATGGCGAG 611
QY 123 GluThrSerGluLeuLeuGluGlnGlnProAlaSerGluAsnValGluGluMetVal 142
DB 612 GAGAACCGAAAT---ATCGAACAGATTCTTCTCAGAAAGCAATTTAGAAAGAGCTTAACA 668
QY 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
DB 669 CAACCCACTGAGTCCAGGCTATGATATTGGATTAAAGAGGTGTTTAAGTTGTTGGC 728
QY 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
DB 729 TTTAAATTCAGTGTGAAAAAGGATAAGACAGAGAAGCCTGCACACTGTCCAGCTACTCACT 788
QY 183 ValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
DB 789 GTGAAGAAAGATGAAGGGGAGGAGCA-----GCAGGGGCTGGCGACCCACGAGAC 839
QY 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
DB 840 CCCAGCCTT-----GGGGCTGGAGAGACGATCTCCAAAGAAAGCGAACCCCAACAATCT 893
QY 223 ThrGluLysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluLeuProLeuGln 242
DB 894 ACAGAGAAACCCGAAGAGACCTTGAACGGTGAGCAAGCCACGACGAGAAATTTCTCCCCCA 953
QY 243 AlaGluSerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLysGlnGlu 262

DB 954 GCCGAATCTGGCCCAACGACGTG---GAGGAATGCCAAAGAGAGGAGAAACAGAA 1010
QY 263 LysGluProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSer 282
DB 1011 AAAGAACCTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTACCAGTGAACAGGATCA 1070
QY 283 SerPheLysLysPhePheThrHisGlyTTPAlaGlyTTPArgLysLysThrSerPheLys 302
DB 1071 ACCTTCAAAAAATTCTTCACTCAAGGTTGGCGCGGTGGCGCAAAAGAACCCAGTTTCAGG 1130
QY 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
DB 1131 AACCCAGAGAGATGAAGTGGAGCTTCCAGAGAGAAAGAAAGAACAGACCCAGAGAAAA 1190
QY 323 ValAspGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
DB 1191 GTAGACACAGAGAACGCGAAAGGCGAGAGTTGCTCTCCGAGAAACTGACCGCCTCCGAG 1250
QY 338 -----GlnGluProAlaGluAspThrAspGlnAlaArgLeuSerAlaAspTyr 353
DB 1251 CAAGCCCAACCCACAGAGCGCGGAGAAAGTCCACAGAGCCCGGTTATCAGCTGAATAT 1310
QY 354 GluLysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu 373
DB 1311 GAGAAAGTTGAGCTGCCCTCAGAGGAGCAAGTCACTGGCTCGCGGAGACCTTCTGAAGAG 1370
QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
DB 1371 AAACCTGCTCCGTTGGCGACAGAGTGTGATGAGAAAAATAGAACTCCACCAAGAGAG 1430
QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGlnGlyGlyGly 412
DB 1431 GTTGTGGCGAAGTCCACGCTCAGCAGCGTGGAGAGAGAACCCGAAAGAGCAG----- 1481
QY 413 GlyLysAlaGluGlyValValValGluGlyThrGlyLysSerLeuProGluLys 432
DB 1482 -----AAACCGGAGGTGGAGAAACACGAGCGGTCTGTGCCAGCTGAAGAA 1526
QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluLeuMetLysSer 452
DB 1527 TTGTTTGGAAATGATGATGAGAACCTCAGGAAGCGCGAACCTGCCAAGAGCTGGTGAAGCTC 1586
QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
DB 1587 AAAGAAACCGTGTGTTTCCGAGAGAGACCTTACACAGGAGAGCTGACCTCAGCTGTATGAG 1646
QY 473 LysThrLeuProLysHisProGluGlyLysValSerGluValMetLeuSerSerGln 492
DB 1647 AAGTGCTGTCCAAACCCCGGAGGCGTGTGAGTGAGTGGAATGTGTCTATCAGAG 1706
QY 493 GluArgLysValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
DB 1707 GAGAGAAATGAGTGCAGGGAAGTCCACTAAAGAAAGCTTTTACCAGCACTGCTTAAAA 1766
QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyGlyGlyGlyGlyGly 532
DB 1767 AAGCTTCTGGAAGAAACAGAAAGGAGAAAGA--GAGAGAGAGACAGAGGAATCAGGG 1823
QY 533 GluTyrGlnHisLysThrGluSerProGluSerAlaAspGluLysGlyGlyGlySer 552
DB 1824 GAGCACTCAGGTTCCAGCCGATCTCCGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGC 1883
QY 553 SerAlaSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
DB 1884 TCTGCTCATCCCTGAGGAGCCCGAGGAGATCAGCTGTCTGGAAAAAGGCTTAGCCGAG 1943
QY 573 AlaProGlnAspGlyGluAlaGluGluGlyThrThrSerAspGlyGluLysLysArgGlu 592
DB 1944 GTGCAGCAGATGGGAGAGCTGAAGAGAGAGAGCTACTTCCGATGAGAGAAAAAGAGAA 2003
QY 593 GlyLysThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
DB 2004 GGTGTCACTCCCTGGGCGATCATTCAAAAAGATGCTGACGCCCAAGAGCGTGTAGACGG 2063

QY 613 ProSerGIuSerAspIysGIuGIuLeuGIuLysValIysSerAlaThrIleuSerSer 632
Db CTTTCGGAAGATGATPAAAGATGAGTGGACAAGGTCAAGAGCCCTTCTGTCTCC 2123
QY 633 ThrAspSerThrValSerGIuMetGIuAspGIuValIleThrValGIuGIuGIuLys 652
Db ACCGAGAGCAGCCCTCTAAATGCAAGAAATGAAAGGAGCGCTGGAGAGCCAAAG 2183
QY 653 ProGIuGIuProLysArgValAspThrSerValSerTrpGIuValIleuIleCysVal 672
Db CCGGAAGAACCAAGCGCAGAGGTGATACCTCAGTATCTTGGGAACCTTTAATTGTGTG 2243
QY 673 GlysSerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGIuGIuGIuProArg 692
Db GGATCATCCCAAGAAAAAGACCAAGAGAGAGGTCTCTTGTGATGAGAGAGCGGACCAAAA 2303
QY 693 ThrLeuGIuLysAspSerHisArgAlaGIuGIuValAserLysAspIysGIuValGIuThr 712
Db GCATAGGGAGAGACCAAGAAAGCTGATGAGGCGGAAAGACCAAGAGAGCGGGACA 2363
QY 713 AspAlaValProAlaSerThrGIuGIuGIuAspGIuAlaGIuLysSerSerProGIu 732
Db GACGGGATCTCTGTCTGTCCCAAGAACATGATCCAGGGCAGGGAAGTTCTCCCGGAG 2423
QY 733 ProAlaGIuSerProSerGIuGIuGIuLysValSerThrTrpGIuSerPheLysArgLeu 752
Db CAAGCTGGAAGCCCTTACCAAGGGGAGGGCGTTTCCACTGGGAGCTATTTAAGGTTA 2483
QY 753 ValThrProArgLysLysSerLysSerLysLeuGIuLysAlaGIuLysP 769
Db GTCACGCCCAAGAAAAAATCAAGTCCAAGCTGGAAGAGAAAAACCAAGACTCCATAGCT 2543
QY 770 ---SerSerValGIuGIuLysSerThrGIuLysLeuGIuLysValGIuLysP 788
Db GGGTCTGGTGAATTCACATTCACAGACTGAACCCGGTAAAGAAATCTCTGGTCT 2603
QY 789 SerLysLysLysPheLysProGIuLysArgLysLysArgAlaAspLysLysGIuGIuLys 808
Db TCAATCAAGAAAGTTTATCTCTGGACAGAGAAAGAAAGCCAGATGGGAACCAAGAACAA 2663
QY 809 AlaThrValGIuAspSerGIuProValGIuLysLeuGIuAspAspProValProAla 828
Db GCCCTGTGGAAGCCAGAGCCAGACAGGGCCCAAGAGATGACTGTGATGTCGGGCC 2723
QY 829 ValAlaProLysSerGIuLysArgAlaValAlaGIuLysArgLysLysPheLysLys 845
Db GTGGTCTCTGTGTGATGATGCTGTGAAGAGGAGAAATGAGGACACAGCAAGCC 2783
QY 846 GIuLysAsnThrGIuLeuProGIuLysLeuGIuLysValIleValIleSerGIuLysLeuSer 865
Db CAAAAGAGCGCAGAGCAGCCCGAGAGGCGCAGCTGAGTGTCCAAAGGAGCTCAAGC 2843
QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGIuThrArgAlaValThrSer 885
Db GAGAGTCAGGTCATATGATGATGAGCAGAGCTGTGCTGAGGGGAGCGAGCTTACCTATT 2903
QY 886 ValGIuGIuLysArgSerProSerTrpLysSerAlaSerValThrGIuProLeuGIuHisThr 905
Db ATTGAAGAAAGGTCTCTCTTGTGATATCTGCTTCAGTGAACAGACTCTTGAACAAGTA 2963
QY 906 AlaGIuGIuAlaMetProProValGIuGIuValThrGIuLysAspLysLysLysLysLys 924
Db GAAGCTGAAGCCGACCTGTTAACTGAGAGGATTTGGAAGAGAAATTAATTGAGAGAA 3023
QY 925 GIuThrProValLeuThrGIuThrLeuProGIuGIuLysAspAlaHisAspAspMetVal 944
Db GAACCCCGCAGGTTACTGACCTTGCAGAGAACAGAGAGCCCGGGGCGCACAGCTC 3083
QY 945 ThrSerGIuValAspPheThrSerGIuAlaValAlaThrAlaThrGIuThrSerGIuAlaLeu 964
Db GTTATGTAGCGGAAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGCAGGGGCCATTG 3143

QY 965 ArgThrGIuGIuValThrGIuAlaSerGIuValGIuGIuThrThrAspMetValSerAla 984
Db GGTTCGGAAGAGAACCGAAGCATCTGCTCTGAAGAGACCAAGAAATGTGTCAACA 3203
QY 985 ValSerGIuLeuThrAspSerProAspThrThrGIuGIuValIleThrProValGIuGIuVal 1004
Db GTCTCCAGTTTACCGACTCCCGACACACAGAGAGGCCACTCCGGTCAAGAGGTG 3263
QY 1005 GIuSerGIuValIleuAspThrGIuGIuGIuLysArgLysGIuThrGIuAlaIleuGIuAla 1024
Db GAAGGTGGCTGATCCGATGATGAGAGAGAGCGGAGCTCAAGAGGTCTCCAGGCA 3323
QY 1025 ValAlaAspLysValIleLysGIuGIuSerGIuValProAlaThr 1038
Db GTGGGAGAAAGGTGAAGAGAAATCCAGCTGCTGGACCGGTGGCCAGAAAGTGTG 3383
QY 1039 ---GIuThrValGIuArgThrGIuSerLysAlaLeuGIuLysValGIuGIuValGIuGIu 1057
Db CTTACGCTGTGCAGAGAGCA-----GAGCGAGAAAGACAGAGAG 3425
QY 1058 AspSerGIuValIleuAlaSerGIuLysGIuLysAspValMetProLysGIuProValGIu 1077
Db CAGCTGAAGCGTCCGGTCTGAGAGAAAGACGATGATGATGTTGAAAGTATGATCTCAG 3485
QY 1078 GIuValGIuAlaGIuHisLysLeuAlaGIuLysArgLysGIuThrGIuAlaThrProGIuSer 1097
Db GAGGCAAAAACCTGAACTTTTACACAAAGGAAAGTGTGGGGAGACCAACCCAGAAAAGC 3545
QY 1098 LeuGIu---ValProGIuValThrAlaAspValAspHis-----ValAlaThrCys 1113
Db TTTGAAAAGCTCTCTCAAGTACAGAGAGCATAGAGTCCAGTGAACCTGTAACTTGT 3605
QY 1114 GIuVal-----IleLysLeuGIuGIuLys---MetGIuGIuAlaValAla 1127
Db CAAGCCGAACCTTACGTGGGTAAATATCACAGAGAGATGATGAACAGAGCTATCCCC 3665
QY 1128 ProGIuSerSerGIuThrLeuThrAspSerGIuThrAsnGIuSerThrProLeuAlaAsp 1147
Db CTTGACTGGTGAAGAACCCCTACAGACAGTGAAGTGAAGAGCAACCCCTGATCCGAC 3725
QY 1148 SerAspThrAlaAspGIuThrGIuLysAspGIuThrIleAspSerGIuAspSerLysAla 1167
Db TTTGACGACCAAGGACCAACCCAGAAAGACAGATGTGGAAATCCATGAGAGAAATAG 3785
QY 1168 ThrAlaAlaValArgLysSerGIuValThrGIuGIuGIuAlaIleThrAlaGIuLysGIu 1187
Db GTGGCATCTGTGATCCAGTCAAGGGGACAGAGAGGAGGAGTCTCGACAGAAAGAG 3845
QY 1188 GIuProSerThrLeuProAsnAsnValProAlaGIuGIuLysLysLysLysLysLys 1207
Db AGGCTCCAGCA---CTTCCAGATTGTGTTCAGGAAAGAACTTAAGAACTTAAGAAAG 3902
QY 1208 ---ArgAspValLeuGIuProThrGIuGIuLysLeuThrAlaAlaValProValLeu 1226
Db ATGGAAGACACTTGAACATACAGATTAAGAGGTGATAGTGAAGAACTGTATCTTCG 3962
QY 1227 AlaLysThrGIuValGIuGIuGIuLysLysLysLysLysLysLysLysLysLysLys 1246
Db TCAAAAGACTGAGGGAGCTCAA-----GAGGCTGACAGTATGCTGATGAGAAACCAA 4016
QY 1247 -----GIuGIuGIuGIuValPheValHisSerGIu-----ProAsnSer 1259
Db GACGTACCATTTTGGAAAGACTTGAAGGGGTCTATAGACAGAGCATTAACGTACGTGG 4076
QY 1260 GIuLysAlaAlaAspValThrLysAspSerGIuValMetGIuValAlaGIuCysGIuGIu 1279
Db GAAAGGTGACACTGAAAGTGGCTTTAAAGTGAAGGACAGAAAGACTGAATGTAAAG 4136
QY 1280 LysGIuSerThrGIuValGIuLysPhe-----LeuSerLeuGIuGIuGIuGIu 1294
Db GATGATGCTCTTGAAGCTGACAGGTACAGCTAAGTCTCTCATCCCGGTGAGAGAGAG 4196
QY 1295 MetGIuThrAspValGIuLysGIuLysArgLysLysLysProGIuGIuValSerGIuGIu 1314

```
Db 4197 ATGGTAGTTCAGTCTGAAGGAGAGAAACAGACAGACGACCCACCTGTGAATGAAGAG 4256
Qy 1315 Gly---GluGlnGluThrAlaAlaProGluHisGluGlyThrTrpGlyProValLeu 1333
Db 4257 AAGCTTGAGCAGCAAAACAGCTGTACCGTATCTCAAGAGGTCAGTAAGCAGCTCCTCCAG 4316
Qy 1334 ThrLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySer--- 1352
Db 4317 ACAGTGAATGTGCCATCATAGATGGGCAAGGAAGATGACAGCTTTTGGAGAGGAGCCCT 4376
Qy 1353 ProSerLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAsp 1372
Db 4377 CTTCCCTGCTAGTCAAGAGGAGGAGTAGTGCACCAAAATTCAGTTTCAAGCTCTGAG 4436
Qy 1373 ThrThrValThrGlnThrAlaGluAlaValGluLysValIle-----GluThrValVal 1390
Db 4437 GCATCACTTCACTTAACAGCGCTGCAGAGGAGGAGAAAGGTCTTGAAGAGAACTGCCAAC 4496
Qy 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
Db 4497 ATTTTAGAAAACAGGTGAACCTTGGAGCCTGCAGGTGCACATTTAGTTCTTGGAGAGAAA 4556
Qy 1411 SerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
Db 4557 TCCTCTGAAAAAATGAAGACTTTGCCGCTCATCCAGGGGGAAGATGCTGTGCCACAGGG 4616
Qy 1431 ProGluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeu 1450
Db 4617 CCGACTGTCCAGCAAAATCGACACCACTGATATCTCTACTTACCAAGAAAGGCTTA 4676
Qy 1451 HisProAspLeuGlnGlyIleLeuSerAlaSerGlnArgGluArgSerGluGluGluAsp 1470
Db 4677 AGTTCCGACCTGGAAGAGAGAGAAAACACATCACTGAAGTGAAGTCAGATGAAGTCGAT 4736
Qy 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
Db 4737 GAGCAGGTTCTGCCAGGAGGTC-----AAAGTGAAGTGTAGCAATTTGAGGATTTA--- 4787
Qy 1491 LysAlaGluProGlu-----IleLeuGluLeuGluSerLysSerAsnLysIleValLeu 1508
Db 4788 -----GAGCCTGAAATGGGATTTTGGNACTTGAGACCAAAAGCAGTAACCTTGTCAA 4841
Qy 1509 AsnValIleGlnThrAlaValAspGlnPheAlaArgThr---GluThrAlaProGluThr 1527
Db 4842 AACATCATCCAGACAGCGCTTGACCACTTTGTACGTACAGAGAAACACAGCCCGAAATG 4901
Qy 1528 HisAlaTrpAspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsn 1547
Db 4902 TTGACGTCTGAGTTACAGACACAGCTCACGTGATATAAAGCTGACACCGAGCGCTGGA 4961
Qy 1548 ArgCysTrpThrLys-----MetLysAspAla-LysMetLysHisProVa 1562
Db 4962 CAGMAACGAGAGAAAGAGGAGGAGGAACTCAGCCCTCGACAGAGATGAACACCAAT 5021
Qy 1562 lProGlnProArgGlu-----AspLeuGlnValLeuThrValLe 1575
Db 5022 ACTTCAGCCAAAGAGGAGTCAAGCTCAACCGCAGTGGGCAAGCACATTCGTATATTTC 5081
Qy 1575 uGluAlaTrpAlaGlnProArgLysCysLeuProArgLeuGlnLeuLysAlaProVal 1594
Db 5082 AAAGACATGATGAAGCCTCAGAAAGACCATGACTCTGTGAGGTAGAAGGTTCCACTG 5139
```

RESULT 5

```
US-09-880-107-3439
; Sequence 3439, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
```

```
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3439
; LENGTH: 6608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U81607
US-09-880-107-3439
```

```
Alignment Scores:
Pred. No.: 1,56e-292 Length: 6608
Score: 4300.50 Matches: 967
Percent Similarity: 68.81% Conservative: 189
Best Local Similarity: 57.56% Mismatches: 408
Query Match: 53.27% Indels: 116
DB: 10 Gaps: 36
```

US-09-902-432-4 (1-1596) x US-09-880-107-3439 (1-6608)

```
Qy 1 MetGlyValaGlySerSerThrGluGlnArgSerProGluGln---ProAlaGlySerAsp 19
Db 192 ATGGGCGCCGGGAGCTCCACCGAGCAGCGAGCCGCGGAGCAGCGCCCGGGGAGGTCC 251
Qy 20 ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer---GlyAla 38
Db 252 ACGCCGCTGAGCGCGAGCCAGCGCGCGGCGGCCCTCGCGCGCGCGCGCCAGACACC 311
Qy 39 AlaGlyAspProAlaAspAla-----AspProAlaThrLysLeuProGlnLysAsnGly 56
Db 312 ACCCGCGACCCCGCCATCGCTGCTCGGACCCCGCCACCAAGCTCTTACAGAGAATGGT 371
Qy 57 GlnLeuSerSerValAsnGlyValAlaGluGlnGlnGlyAspValHisValGlnGluAsn 76
Db 372 CAGCTGTCCACCATCAATGGCGTAGCTGACCAAGATGAGCTCAGCTCCAGGAGGTGAC 431
Qy 77 GlnGluGlyGln-----GluGlu 82
Db 432 CTAAATGGCCAGAAAGAGAGCCCTGAACGGTCAAGGAGCCCTAAACAGCCAGGAGAGAA 491
Qy 83 GluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArg 102
Db 492 GAAGTCATTTGTCACGGAGGTGGACAGAGAGACTCTGAAGATGTGAGCGAAAGAGACTCC 551
Qy 103 ValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGlnGlu 122
Db 552 GATAAAGATGCTACTAAGTCAGCGGTGTTTCACGACATCACAGATGATGGCCAGAG 611
Qy 123 GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnValGluGluMetVal 142
Db 612 GAGAACCGAAAT--ATCGAACAGATTCTTCTTCAGAAAGCAATTTAGAAGAGCTAACA 668
Qy 143 GlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGly 162
Db 669 CAACCCACTGAGTCCCGAGGCTAATGATATTGGATTTTAAAGAGGTGTTTAAAGTTTGTTC 728
Qy 163 PheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr 182
Db 729 TTTAAATTTCACTGTGAAAAAGGATAAGACAGAGAGCTGCACACTGTCCAGTACTCACT 788
Qy 183 ValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
Db 789 GTGAAGAAAGATGAAGGGGAGGAGCA-----GCAGGGGCTGGCGACACCAGGAC 839
Qy 203 ProSerValGluThrAlaValGlnGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
```

Db 840 CCCAGCCTT-----GGGGCTGGAGAGACGATCCAAAGAAAGAAAGCCAAACATCT 893
 QY 223 ThrGluysgIngluGlyThrLeuysgIngluInserSerThrGluIleProleuGln 242
 Db 894 ACAAGAAAGCCGGAAGAGACCTTGAAGCGTGAGCAAGCCAGCAAGAAATTTCTCCCCA 953
 QY 243 AlaGluSerAspGlnAlaIleGluGluAlaIleAspGluGluGluGluGln 262
 Db 954 GCCGAATCTGGCCGAAGAGTG---GAGGAATGCAAGAGAAAGAGAAAGAAAGAA 1010
 QY 263 LysGluProThrLysSerProGluSerProSerProValAsnSerGluThrThrSer 282
 Db 1011 AAAGAACCTAGAAAGTCTGCAGAAATCTCCGACTAGTCCCTGACCAAGTAAACAGATCA 1070
 QY 283 SerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSerPheLys 302
 Db 1071 ACCTTCAAAAAATTTCTTCACTCAAGTTGGCGCGGTGGCGCAAAAAAGCCAGTTTCAGG 1130
 QY 303 LysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGlnGluAlaGluLys 322
 Db 1131 AAGCCGAAGAGAGATGAAGTGAAAGTCTCAGAGAAAGAAAGAAAGAAAGCCAGAAAAA 1190
 QY 323 ValAspGluGluGluLysGluLysThrGluProAlaSerGluGlu----- 337
 Db 1191 GTAGACACAGAAAGAGACGAAAGGACAGAGTTGCTCCGAGAAACTGACCCGCTCCGAG 1250
 QY 338 -----GIngluProAlaGluAspThrAspGlnAlaArgLysSerAlaAspTyr 353
 Db 1251 CAAGCCACCCACAGAGAGCGGCGCAAGAAAGTGCCCAAGAGCCCGGTTATCACTGAATAT 1310
 QY 354 GluLysValGluLeuProleuGluAspGlnValGlyAspLeuGluLysSerGluGlu 373
 Db 1311 GAGAAAGTTGAGCTCCCTCAGAGAGCAAGTCAAGTGCTCCGAGAACCTTCTGAGAGG 1370
 QY 374 LysCysAlaProLeuAlaThrGluValPheAspGluLysMetGluAlaHisGln---Glu 392
 Db 1371 AAACCTGCTCGCTGGCGAGCAAGAGTGTGATGAGAAATATGAAAGTCCACCAAGAAAG 1430
 QY 393 ValValAlaGluValHisValSerThrValGluLysThrGluGluGluGlnGlyGly 412
 Db 1431 GTTGGGCGGAAGTCACTGACGACCGTGGAGAGAAAGCAAGAGACAG----- 1481
 QY 413 GlyIleAlaGluGlyValValValGluGlyThrGlyLysLeuProProGluLys 432
 Db 1482 -----AAAAAGAGGTGGAAGAAACGACGAGGTCTGTGCTGACCTGAGAA 1526
 QY 433 LeuAlaGluProGlnGluValProGlnGluAlaGluProAlaGluGluMetLysSer 452
 Db 1527 TTGGTTGGAAATGATGACAGAACCTCAGAAAGCCGAAGCTGCGCAAGAGGCTGGTGAAGCTC 1586
 QY 453 ArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGlu 472
 Db 1587 AAAGAAACGTGTGTTCCGGAAGAGACCTTACACAGAGGAGCTGACTGCTGATGAG 1646
 QY 473 LysThrLeuProLysHisProGluGlyIleValSerGluValGluMetLeuSerSerGln 492
 Db 1647 AAGGTGCTGTCCAAACCCGCCAAGGCGTTGTGATGAGTGAAGTGAATGCTGTATCATCAG 1706
 QY 493 GluArgGluIleValGlnGlySerProLeuLysLysLeuPheSerSerSerGlyLeuLys 512
 Db 1707 GAGAAATATGAGGTGCAGGAGAGTCCATTAAGAACCTTTTACACGCACTGGCTTAAAA 1766
 QY 513 LysLeuSerGlyLysLysGlnLysGlyLysArgGlyGlyGlyLysAspGluGluProGly 532
 Db 1767 AAGCTTTCTGGAAGAAAGAAAGAAAGAAAG---GAGAGAGAGAGACAGAGATAGAGG 1823
 QY 533 GluTyrGlnHisIleHisThrGluSerProGluSerAlaAspGluGluLysGlyLysSer 552
 Db 1824 GAGCACACTCAGGTTCACCGCATTTCTCCGACAGCCAGAGAGAGAGAAAGGCGAGAGAC 1883
 QY 553 SerAlaSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyLysProleuGlu 572
 Db 1884 TCTGCTCATCTCCCTGAGAGGCCGAGAGAGATCACTGTCTGGAAGAAAGGCGTTAGCCGAG 1943

QY 573 AlaProGlnAspGlyValAlaGluGluGlyThrThrSerAspGlyLysLysArgGlu 592
 Db 1944 GTGACAGCGATGGGGAGAGCTGAAGAAAGAGAGCTACTTCCGATGAGAGAAAAAGAGAAA 2003
 QY 593 GlyIleThrProTrpAlaSerPheLysLysMetValThrProLysLysArgValArgArg 612
 Db 2004 GGATCACTCCCTGGGCACTCATTCAAAAGATGGAGAGCCCAAGAAAGCGTTTATGAGCGG 2063
 QY 613 ProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThrLeuSerSer 632
 Db 2064 CTTTCGAAATATGATTAAGAGATGAGCTGAGCAAGGTCAAGACGCTTACTTGTCTCC 2123
 QY 633 ThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGluGluLys 652
 Db 2124 ACCGAGACACAGCGCTTGAATGCAAGAAATGAAGAAAGGACGCTGGAAGAGCCAAAG 2183
 QY 653 ProGluGluProLysLysArgValAspThrSerValSerTrpGluAlaLeuIleCysVal 672
 Db 2184 CCGAAGAAACCAAGCGCAAGGTGATCTTCAGTATCTTGGAAAGCTTTAATTTGTGTG 2243
 QY 673 GlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAspGluGlyProArg 692
 Db 2244 GGATCATCCAAAGAAAGAGCAAGAGAGATCTCTTCTGATGAGAAAGGGGAGCAAAA 2303
 QY 693 ThrLeuGlyLysAspSerHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThr 712
 Db 2304 GCATGTGAGAGAGACCAACCAAGAAAGCTGATGAGCGCGGAAAGAAAGACGAGGAGACA 2363
 QY 713 AspAlaValProAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProLys 732
 Db 2364 GACGGGATCTCTTGCTGTTCCTCAAGAAACATATCCAGGCGAGGAAAGTTCTCCCGGAG 2423
 QY 733 ProAlaGlySerProSerGluGlyGlyValSerThrTrpGluSerPheLysArgLeu 752
 Db 2424 CAAGCTGAAACCTTACCGGAAGGGAGGCGTTTCCACTGGAGTCAATTAAAAAGTTA 2483
 QY 753 ValThrProArgValLysSerLysSerLysLeuGluGluValAlaGluAsp----- 769
 Db 2484 GTCAAGCCCAAGAAAAAATCAAGTCCAAAGCTGGAAGAGAAAGCAAGACCTCATAGCT 2543
 QY 770 ---SerSerValGluGlnLeuSerThrGluIleGluProSerArgGluGluSerThrVal 788
 Db 2544 GGGTGTGTGTAAGACATTCCATCCAGACACTGAACCGGTAAAGAAATCTCGGGTCTC 2603
 QY 789 SerIleLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGln 808
 Db 2604 TCAATCAAGAAAGTTATTCTCTGAGAGAGAAAGAAAGCCAGATGGGAAACAAAGACAA 2663
 QY 809 AlaThrValGluAspSerGlyProValGluIleGlnGluAspAspProAsnValProAla 828
 Db 2664 GCGCCCTGTGAAGACGACGGGCCAAACAGGGCCAAACGAAGATGACTGTATGATCCGGCC 2723
 QY 829 ValValProLeuSerGluTyrAsnAlaValAlaGluArgGlyLysMetGlu-----Ala 845
 Db 2724 GTGTGCTCTGTCTGAATGATGATCTTGAAGAGGAGAAATGAGACACAGCAAGCC 2783
 QY 846 GlnGluAsnThrGluLeuProGlnLeuGluAlaValAlaValSerGluGluLeuLeuSer 865
 Db 2784 CAAAAGCGCGCAGAGACACCCGAGCAGAAAGCAGCACTGAGAGTCTTCAAGAGAGCTCAC 2843
 QY 866 LysThrLeuValHisThrValSerValAlaValIleAspGlyThrArgAlaValThrSer 885
 Db 2844 GAGATCAGGTTCATATATGATGAGCAGCACTGTGCTGACGAGGAGGAGCTTACATT 2903
 QY 886 ValGluGluArgSerProSerTrpIleSerAlaSerValThrGluProleuGlnHisThr 905
 Db 2904 ATTGAAGAAAGGTCTCTTCTTGTGATATCTGCTTCAAGAGAAACCTTGAAGAACAGTA 2963
 QY 906 AlaGlyGluAlaMetProProValGluGluValThrGluLysAspIleIleAla---Glu 924
 Db 2964 GAACTGAAGCCGCACTGTAACTGAGAGAGTATTGAAAGAGATATTGCAAGAGAA 3023

Qy	925	GluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHisAspAspMetVal	944
Db	3024	GAACCCCCACGGTTACTGAACCTCTGCAGAGAACACAGAGAGCCCGGGCGACACGGTC	3083
Qy	945	ThrSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeu	964
Db	3084	GTTAGTGAGCGGAATTGACCCCGAAGCTGTGACAGCTGCAGAAACTGTGACGGCCATTG	3143
Qy	965	ArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAla	984
Db	3144	GGTTCCGAAGAGAAACCGAAGCATCTGCTGCTGAAGAGACCACAGAAATGCTGTACAGCA	3203
Qy	985	ValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValGlnGluVal	1004
Db	3204	GTCTCCAGTTAAACCACTCCACAGACACACACAGAGGAGGCCACTCCGCTGCAGAGGTG	3263
Qy	1005	GluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAla	1024
Db	3264	GAAGTGGCTGACTGACATAGAAAGCAAGAGAGCGGACTCAAGAGGTCTCCACAGCA	3323
Qy	1025	ValAlaAspLysValLysGluGluSerGlnValProAlaThr-----	1038
Db	3324	GTGGCAGAAAGATGAAGAGGAATCCACAGTCCCTGGCACCGGTGGCCACAGAGATGTG	3383
Qy	1039	---GlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluValGluGlu	1057
Db	3384	CTTCAGCCTGTGCAGAGACA-----GAGGCAGAAAGACCACAGAGAG	3425
Qy	1058	AspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProValGln	1077
Db	3426	CAGGCTGAAGCGTGGGTCTGAAGAAGAGACGCGATGTAGTGTGTGAATAGTAGTCTCAG	3485
Qy	1078	GluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrProGluSer	1097
Db	3486	GAGCAAAACTGAGCCTTTTACACAGGGAAGTGTGTGGGCGAGACCAACCCCAAGAACG	3545
Qy	1098	LeuGlu---ValProGluValThrAlaAspValAspHis-----ValAlaThrCys	1113
Db	3546	TTTGAAAAAGCTCTCAAGTCAACAGAGACATAGAGTCCAGTGAGCTTGTAAACCACTGT	3605
Qy	1114	GlnVal-----lleyLysLeuGlnGlnLeu---MetGluGlnAlaValAla	1127
Db	3606	CAAGCGAAACCTTAGCTGGGTGAAATCACAAGGAGATGGTGTGAACACAGGCTATCCCC	3665
Qy	1128	ProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAsp	1147
Db	3666	CCTGACTCGGTGGAACCCCTACAGACACTGAGACTGATGAAGCACCCCGCTGACCCAG	3725
Qy	1148	SerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLysAla	1167
Db	3726	TTTGACGCACAGGCACAACCCAGAAAGACAGAGATTGTGGAAATCCATGAGGAGAAATGAG	3785
Qy	1168	ThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGlu	1187
Db	3786	GTCGCATCTCGTACCAGTACAGGGGCACACAGAGCAGAGGCGAGTTCTCTGCACAGAAAG	3845
Qy	1188	GluProSerThrLeuProAsnValProAlaGlnGluGluHisGlyGluGluProGly	1207
Db	3846	AGGCCTCCAGCA--CCTTCCAGTTTGTGTCTCCAGGAAGAAACTAAGAAACAATCAAG	3902
Qy	1208	---ArgAspValLeuGluProThrGlnGlnLeuThrAlaAlaAlaValProValLeu	1226
Db	3903	ATGGAACACACTCTAGAGCATACAGATAAAGAGGTGTGAGTGGAACTGTATCTCATTTCTG	3962
Qy	1227	AlaLysThrGluValGlyGlnGluGlyGluValAlaAspTrpLeuAspGlyGlyLysValLys	1246
Db	3963	TCAAAGACTGAGGGGACTCAA-----GAGGCTGCCACGATATGCTGATGAGAAACCCAAA	4016
Qy	1247	-----GluGluGlnGlnValPheValHisSerGly-----ProAsnSer	1259
Db	4017	GACGTACCATTTTCGAAGGACTTAGGGGCTCTATAGACACAGGCATAACAGTCAGTCGG	4076
Qy	1260	GlnLysAlaAlaAspValThrTrpAspSerGluValMetGluValAlaGluCysGlnGlu	1279

RESULT 6
US-09-74

Db 1681 GAAAATGTTGAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAA 1740
QY 562 GluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGlu 581
Db 1741 GAAAGTGTAGCTGAAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCT 1800
QY 582 GlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAlaSerPheLys 601
Db 1801 CCAACTGTT-----GAAGAAATGTAGCTCCA-----AGTGTGTGA 1836
QY 602 LysMetValThrProLysLysArgValArgProSerGluSerAspLysGluGluGlu 621
Db 1837 GAAAGTGTGCTCCAAGT-----GTTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGT 1890
QY 622 Leu-----GluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet 640
Db 1891 GTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAAT 1950
QY 641 GlnAspGluVal-----LysThrValGlyGluGluGlnLysProGluGluProLysArg 658
Db 1951 GTTGAAGAAATCGTAGCTCCAAGTGTGGAAGAAATCGTAGCTCCA-----1995
QY 659 ArgValAspThrSerValSerTrpGluAlaLeuIleCysValGlySerSerLysLysArg 678
Db 1996 -----ACTGTTGAAGAAATGTAGCTCCAAGTGTGGAAGAAAGTGTG 2037
QY 679 AlaArgLysAlaSerSerAspAspGluGlyGlyProArgThrLeuGlyGlyAspSer 698
Db 2038 GCTCCAAGTGTGGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTG-----2085
QY 699 HisArgAlaGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer 718
Db 2086 -----GCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGT 2130
QY 719 ThrGlnGluGlnAspGlnAlaGlnGlySerSerProGluProAlaGlySerProSer 738
Db 2131 GTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAAGT 2190
QY 739 GluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrPro-----755
Db 2191 GTTGAAGAAATCGTAGCT-----CCAACTGTTGAAGAAATGTTAGCTCCAAGTGTGTA 2244
QY 756 -----ArgLysLysSerLysSerLysLeuGluGluLysAlaGluAspSerSerVal 772
Db 2245 GAAAGTGTGCTCCAAGTGTGGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTAGCT 2304
QY 773 GluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLys 792
Db 2305 GAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGT-----2346
QY 793 PheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGlnAla-----Thr 810
Db 2347 -----GTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAAGT 2382
QY 811 ValGluAspSerGlyProValGluIleAsnGluAspAspProAsnValProAlaValVal 830
Db 2383 GTTGAAGAAATGTTAGCTCCAAGTGTGGAAGAA-----AGTGTAGCT 2424
QY 831 ProLeuSerGluTrpAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGlu 850
Db 2425 CCAAGTGTGGAAGAAAGTGTGCTGAAAACGTTGCCAACAAATTTATCAGACAACTCTT---2481
QY 851 LeuProGlnLeuLeuGlyAlaValTyrValSerGluGluLeuSerLysThrLeuValHis 870
Db 2482 TTAAGTAATTTATTAGTGTGATC---GAAACTGAGGAATAAAGACAGATATATTAAT 2538
QY 871 ThrValSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSer 890
Db 2539 GAGATAGAA-----GAAGTAAAAAGAAAT---2562
QY 891 ProSerTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMet 910

Db 2563 -----GTAGTCACCACCACTACTAGAAAACGTAGAAAACACTACAGCTGAAAGTGTA 2613
QY 911 ProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProValLeuThr 930
Db 2614 ACTACTTTTAGTAACATATTAGAGGAG-----ATACAAGAAAATACTATTACTAAT 2664
QY 931 GlnThrLeuProGluGly---LysAspAlaHisAspAspMetValThrSerGluValAsp 949
Db 2665 GATACTATTAGAGGAAAATTTAGAAAGAACCTCCACGAAAATGTATTAAAGTCGCGCTTAGAA 2724
QY 950 PheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluVal 969
Db 2725 AATACCCAAAGTCAGAGGAAAAGAAAGTAATAGATGTATTT-----GAAGAAGTA 2778
QY 970 ThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAlaValSerGln-----987
Db 2779 AAAGAA-----GAGGTCGCTACCACCTTTAATAGAAACTGTGGAACACGCGAGAA 2826
QY 988 -----LeuThrAspSerProAspThrThrGluGluAlaThrProVal 1001
Db 2827 GAAAAGAGCGCAATAACAATTACGSAATAATTTGAAAATTTTAGAAGAA-----2874
QY 1002 GlnGluValGluSerGlyValLeuAspThrGluGluGluArgGlnThrGlnAlaIle 1021
Db 2875 AATGCGAGTAGAAGTAATGAAAATGTTGCAGAGAAATTTAGAGAAATTAACAGAACTGTA 2934
QY 1022 LeuGlnAlaValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrVal 1041
Db 2935 TTTAATACTGTATTAGATAAAGTAGAGAAACAGTAGTAAATTTAGCGAGAAAGTTTAGAA 2994
QY 1042 GlnArgThrGlySerLysAla-----LeuGluLysValGluGluValGlu 1056
Db 2995 AACAAATGAAATGGAATAAGCAATTTTTTAGTGAAATATTTGATAATGTAAAAGGAATACAA 3054
QY 1057 GluAsp-----SerGluValLeuAlaSerGlu 1065
Db 3055 GAAAATTTAATAACAGGTATGTTCCGAAGTATAGAAACCAGTATAGTAATCCAATCAGAA 3114
QY 1066 LysGluLysAspVal-----Met 1071
Db 3115 GAAAAGTTGATTTGAATGAAAATGTTGTTAGTTCCGATTTAGATAATATAGAAAATATG 3174
QY 1072 ProLysGly-----ProValGlnGluAla 1079
Db 3175 AAAGAAGGTTTATTAATAATAATAGAAAATATTTCAAGTACTGAAGTGTGTCAAGAACT 3234
QY 1080 GlyAlaGluHisLeuAlaGlnGly-----1087
Db 3235 GTAACTGAAACATGTAGAACAAAATGTATATGTGGATGTTGATGTTCTGCTATGAAAGAT 3294
QY 1088 -----SerGluThrGlyGlnAlaThrProGluSerLeuGluValPro 1101
Db 3295 CAATTTTATAGGAATATAAATGAGCGAGGCGGTTGAAAAGAAATGTTTTTAATTTGGAA 3354
QY 1102 GluValThrAlaAspValAspHisValAlaThrCysGlnValIleLys-----Leu 1118
Db 3355 GATGATTTTAAAGAGTGAAGAGTGATGTAATTAATCTGTAGAGAAATTAAGGATGAACCGGT 3414
QY 1119 GlnGlnLeuMetClnGln-----AlaValAlaProGluSerSerGluThrLeuThr 1135
Db 3415 CAAAAGAGGTAGAGAAAGAACTGTTGATTTATTGTAAGAAATGGAAGAAATATTGTA 3474
QY 1136 AspSerGluThrAsnGlySerThrProLeuAlaAspSerAspThrAlaAspGlyThrGln 1155
Db 3475 GATGATTAGAGGAGAAAAGAGATTTAAACAGACAAG---ATGATAGATGACGTAGAA 3531
QY 1156 GlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaValArgGlnSerGln 1175
Db 3532 GAATCCATAGAAATATCTTCAGATCTTAAGAGAAAGAACTGAATCTATTAAAGATAAA---3588
QY 1176 ValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsn 1195
Db 3589 -----GAAAAGAGATGTTTCACTAGTTGTTGAAGAA-----3618

Db	866	TTTAATAGTTAGTAAAA	-----AGTGTTCACAAAGAACAAACAC	907
Qy	174	LysSerAspThrValGlnLeuLeuLeuThrValIysLysAspGluGlyGlnGlyAlaGluAla	193	
Db	908	AATGTTCAAGAAAAAGTTGAAGAAAGTGTGAAGAAATGACGAAGAAGTGTAGAGAA	967	
Qy	194	SerVal-----GlyAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGlu	211	
Db	968	AATGTGAAGAAAAATGTAGAAGAAATACGACGGAAGTGTAGCCTCAAGTGTGAAGAA	1027	
Qy	212	Ser-----AlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThr	229	
Db	1028	AGTATAGCTTCAAGTGTGATGAAGATAGATAGAAATCAAGTATTGAAGAAATGTAGCTCCA	1087	
Qy	230	LeuLysGlnGlu-----	233	
Db	1088	ACTGTTGAAGAAATCGTAGCTCCAAGTGTGTAGAAAGTGTGGCTCCAAGTGTGAAGAA	1147	
Qy	234	-----GlnSerSerThrGluLeuProLeuGlnAlaGluSerAspGlnAlaAlaGluGlu	251	
Db	1148	AGTGTGAAGAAAAATGTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAA	1207	
Qy	252	GluAlaLysAspGluGlyGlnGluLysGlnGluLysGluProThrLysSerProGluSer	271	
Db	1208	AATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAA	1267	
Qy	272	ProSerSerProValAsnSerGluThrThrSerSerPheLysLysPhePheThrHisGly	291	
Db	1268	ATCGTAGCTCCAAGTGTGAAGAAATCGTAGCTCCACTGTTGAAGAAATGTTAGCTTCCA	1327	
Qy	292	TrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAspAspLeuGlu---	310	
Db	1328	AGTGTGTAGAAAAAGTGTGGCTCCAAGTGTGTGAAGAAAGTGTAGAGAAAAATGTTGAAGAA	1387	
Qy	311	---ThrAlaGluLysArgLysGlnGlnGluAlaGluLysValAspGluGluGluLysGlu	329	
Db	1388	AGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAA	1447	
Qy	330	LysThrGluProAla-----SerGluGluGlnGluProAlaGluAspThrAspGln	346	
Db	1448	AATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAA	1507	
Qy	347	AlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeuGluAspGlnValGlyAsp	366	
Db	1508	ATC---GTAGCTCCAAGTGTGAAGAAATCGTAGCTCCAAGTGTGAAGAAAT---	1558	
Qy	367	LeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThrGluValPheAspGluLys	386	
Db	1559	GTAGCTCCAAGTGTGTAGAAAGTGTGGCTCCAAGTGTGTGAAGAAAGTGTGAAGAAAT	1618	
Qy	387	MetGluAlaHisGlnGlnValValAlaGluValHisValSerThrValGluLysThrGlu	406	
Db	1619	GTGTGA-----GAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGA	1672	
Qy	407	GluGluGlnGlyGlyGlyGluAlaGluGlyGlyValValValGluGlyThrGlyGlu	426	
Db	1673	GAAAGT-----GTAGCTGAAAAATGTTGAAGAA	1699	
Qy	427	SerLeuProProGluLysLeuAlaGluProGlnGluValProGlnGluAlaGluProAla	446	
Db	1700	AGTGTA-----GCTGAAAAATGTT	1717	
Qy	447	GluGluLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThrGlnLeuThr	466	
Db	1718	GAAAGAAAGTGTAGCTGAA	1735	
Qy	467	AspLeuSerProGluGluLysThrLeuProLysHisProGluGlyIleValSerGluVal	486	
Db	1736	-----AATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTT	1789	
Qy	487	GluMetLeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLysLysLeuPhe	506	
Db	1790	GAAAGAAATCGTAGCTCCAAGTGTGTGAAGAAATCGTAGCTCCAAGTGTGTGAAGAAATGTA	1849	

QY	507	SerSerSerGlyLeuLysLysLeuSerGlyLysGlnLysGlyLysArgGlyGly	536
DB	1850	GCCTCAAGTGTGTGAAGAAAGTGTGGCT	1876
QY	527	GlyAspGluGluProGlyGlnThrGlnHisIleHisThrGluSerProGluSerAlaAsp	546
DB	1877	-----CCAAAGTGTGAAGAAAGTGTAGAAGAAATGTTGAAGAAAGTGTAGCT	1924
QY	547	GluGlnLysGlyLeuSerSerAlaSerSerProGluGluPro	561
DB	1925	GAATAATGTTGAAGAAAGTGTACTGAAATGTTGAAGAAAGTGTACTGAAATGTTGA	1984
QY	562	GluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGlu	581
DB	1985	GAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCT	2044
QY	582	GlyThrThrSerAspGlyGluLysLysArgGluGlyIleThrProTrpAlaSerPheLys	601
DB	2045	CCAACCTGTT-----GAAGAAATGTTAGCTCCA-----ACTGTTGTA	2080
QY	602	LysMetValThrProLysLysArgValArgArgProSerGluSerAspLysGluGluGlu	621
DB	2081	GAAGTGTGGCTCCAAGT-----GTTGAAGAAAGTGTAGAAGAAATGTTGAAGAAAGT	2134
QY	622	Leu---GluLysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMet	640
DB	2135	GTAGCTGAAATGTTCAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAAT	2194
QY	641	GlnAspGluVal-----LysThrValGlyGluGluGlnLysProGluGluProLysArg	658
DB	2195	GTTGAAGAAATCGTAGCTCCAAGTGTGTAAGAAATCGTAGCTCCA-----	2239
QY	659	ArgValAspThrSerValSerTrpGluAlaLeuLysCysValGlySerSerLysLysArg	678
DB	2240	-----ACTGTTGAAGAAATGTTAGCTCCAAGTGTGTAGAAAGTGTG	2281
QY	679	AlaArgLysAlaSerSerSerAspGluGlyGlyProArgThrLeuGlyGlyAspSer	698
DB	2282	GCTCCAAGTGTGAAGAAAGTGTAGAAGAAATGTTGAAGAAAGTGTGA-----	2329
QY	699	HisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSer	718
DB	2330	-----GCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGT	2374
QY	719	ThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSer	738
DB	2375	GTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAAT	2434
QY	739	GluGlyGluGlyValSerThrTrpGluSerPheLysArgLeuValThrPro	755
DB	2435	GTTGAAGAAATCGTAGCT-----CCAAGTGTGAAGAAATGTTAGCTCCAAGTGTGTA	2488
QY	756	-----ArgLysLysSerLysSerLysLeuGluGluLysAlaGluAspSerVal	772
DB	2489	GAAGTGTGGCTCCAAGTGTGAAGAAAGTGTAGAAGAAATGTTGAAGAAAGTGTAGCT	2548
QY	773	GluGlnLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysLys	792
DB	2549	GAATAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGT-----	2590
QY	793	PheIleProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGlnAla	810
DB	2591	-----GTAGCTGAAATGTTGAAGAAAGTGTAGCTCCAAGTGTAGCTCCAAT	2626
QY	811	ValGluAspSerGlyProValGluIleAsnGluAspProAsnValProAlaValVal	830
DB	2627	GTTGAAGAAATGTAGCTCCAAGTGTGAAGAA-----ACTGTTAGCT	2668
QY	831	ProLeuSerGlyThrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGlu	850
DB	2669	CCAAGTGTGAAGAAAGTGTGCTGAAACGTTGCAACAAATTTATACAGCAATCTT---	2725


```
Db 4730 GAAGAAGTATTATTATTAAGAGAGGTTTAAAGAGAACCAACCAAGAAAAATAACAAAAAAG 4789
Qy 1415 GlycylHisTrpThrLeuGlnHisAlaGlu-----AspThrValProLeuGlyProGlu 1432
Db 4790 AAAGTAAAGTTTGTATATTAGGATTAAGGATTAAGCAACCAAGATGAAATA---GTAGAAGTTGAA 4846
Qy 1433 SerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeuHisPro 1452
Db 4847 ATGAAGATGAAGATATA-----GAAGAAGATGTAGAAGAA 4882
Qy 1453 AspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAspLysPro 1472
Db 4883 GATATAGAACAAGATATAGAAGATAAAGTTGAAGATATAGATATAGATAGAA 4942
Qy 1473 AspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeuLysAla 1492
Db 4943 GATATAGTGAAGACAAAGAT-----GAAGTTATAGATTTAATAGTCCAAAAA 4990
Qy 1493 GluProGluIleLeuGluLeuGluSerLysSerAsnLysIle 1506
Db 4991 GAGAAACGCATTGAAAGGTTTAAAGCGAAAGAAAAATA 5032
```

RESULT 8

```
US-09-864-761-17726
; Sequence 17726, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

```
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 17726
; LENGTH: 7997
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL034555.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.3
; OTHER INFORMATION: EXPRESSED IN BT100, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN H474, SIGNAL = 12
; OTHER INFORMATION: EST HUMAN HIT: AUL17052.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P08640, EVALUE 4.00e-10
; OTHER INFORMATION: NT HIT: gi11427711, EVALUE 0.00e+00
US-09-864-761-17726
```

Alignment Scores:

Pred. No.:	3 69e-21	Length:	7997
Score:	431.50	Matches:	373
Percent Similarity:	34.27%	Conservative:	271
Best Local Similarity:	19.85%	Mismatches:	759
Query Match:	5.34%	Indels:	476
DB:	10	Gaps:	83

US-09-902-432-4 (1-1596) x US-09-864-761-17726 (1-7997)

Qy	31	ProAlaAlaGluAlaSerGlyAlaAlaGly-AspProAlaAspAlaAspProAla	48
Db	886	CCTGGCAGCAAACTTCAGGTCTCTCAGCGAGCGCTGCAAAATCTGACTGTCTAACT	945
Qy	49	-----ThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAsnGlyVa	64
Db	946	GGAATCAGTTAGAAATGAAAGTACCAAGGAAAGGGGCTTTCAAGGCCATCTTGAAGTGT	1005
Qy	64	lAlaGluGlnGlyAspVal-----HisValGlnGluGlnGlnGluGlyG1	80
Db	1006	GGAGAAGGAGGCGGCTTAAAGCCAGGAGCACCTCAGCCTGAG---CAGCCTGAGAG	1062
Qy	80	nGluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLy	100
Db	1063	TGGGGTAAGTGTGTGAT-----	1081
Qy	100	sAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspG1	120
Db	1082	-----CTGGAGAAGCTGGAAGCCAGGAAAGGCGCTTGCAGATTTCAATTTAAAGC	1134
Qy	120	yGlnGluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsn---AsnValG1	139
Db	1135	AGAAAAGCAAAACCCAGAGGTCAAGAAAGAGCAGTCCAGAGATGGAGGATGCTCGGTGCT	1194
Qy	139	uGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPhe--	158
Db	1195	TTCAAAAAGCAGCCT-----GACGTGCTCTTAGAGAGGTCTATTCT	1236
Qy	159	-----LysPheValGlyPheLysPheThrValLysLysAs	170
Db	1237	GCTGAGGGAAGGAGAGGCTGAAAGAAAGCCTGTGAGGAAAGAAATTTTAAAGAGATC	1296
Qy	170	pLysAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGlyG1	190
Db	1297	TAAAAAAATCAAACTGGACAGACTTAATCTGTTCAGCCCAAGACTGTGAGGAGCT	1356
Qy	190	yAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSerValGluThrAlaValG1	210
Db	1357	TGCCAGTATTCTGTGGGTCTGGCTCAAGGCCAGCTCAGACCTCAAGAGACTGGG	1416
Qy	210	yGluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThrLe	230

```

Db 1417 AGAAGTACGAGT-----GAATCTGGGAAATCAAGAA----- 1450
QY 230 ULySgInglInserSerThrGluLePro-----LeuGlnAl 243
Db 1451 -----GTCCATCAAAAGAAAGCCCATCTCCCAAAACACAGCTCAACAGCTGAGGT 1503
QY 243 AGIuserSerGlnAlaIaIaGluGlu-----AlaIaSerGluGluGluGluGlu 261
Db 1504 ATTAGATGATTCAGAGCAACAGAGAGAGAGACGTTAGGAAAAAATATTGAGTCTTGTA 1563
QY 261 nGluYsGluProThrLysSerProGluSerProSerSerProValAsnSerGluThr-- 280
Db 1564 TGAAACACCTGAACGTAATCAAGGCCAAGAGAAATCACTTCACTAAATATCTGAAGAAA 1623
QY 281 -----ThiSerSerPheLysLysPhePheThrHisGlyTTrpAl 293
Db 1624 AATTGGCATTGACATCGATCAACGAGAGTTACGAAAAAATAAGAACAGAGTCTGAG 1683
QY 293 aGlyTTrpAlGlyLysLysThrSerPheLysSerLys-----GluAspAs 308
Db 1684 GAAACAGCAGATGAGAAATGCAAAATACCAAGTCTGAAAGTTGGCATCTCTAAAAAGAA 1743
QY 308 PLeuGluThrAlaGluLysArg----- 315
Db 1744 TGTAGATGAATGTGAAGAGCGTAGCCTCTTCAACAGTAGGCAAAACCCCTCAAGATGT 1803
QY 316 -----LysGluGlnGluAlaGluLysValAsp----- 324
Db 1804 CACTGATGACTCTCCCTCCTAGCAAAAGAAAGAGATGATCATGTCATTTTGATATCTG 1863
QY 325 -----GluGluGluLysGluLysThrGluProAlaSerGluGlnGluProAl 341
Db 1864 CACCAAGCGAGAAACGGAATTAACAGAAATTCACGCCCAATACAGGAGAAATTCGAAAGAAC 1923
QY 341 aGluSerThrArgGlnAlaIaIaGluSerAlaAspGlyLysValGluLeuProLeuG 361
Db 1924 TGGTGTTCTCCAGTGTCCGACATGTTCTTCCATGAAGAT-----GA 1968
QY 361 uAspGlnValGlyAspLeuGluAlaSerSer-----GluGluLys 374
Db 1969 GGATCCCATAGGTCTCCCTAGGCTACTCTCAGTAAAGAGGTCTCTAAAGATGATGAAA 2028
QY 374 sCysAlaPro-----LeuAlaThrGluValPheAs 384
Db 2029 ACTCTCTCCCTTCTTACATTAACAGTCAAGGAGAGTCTTTAAATTAATCTTATGA 2088
QY 384 pGluLysMetGluAlaHisGlnGluValAlaIaGluValHisValSerThrValGluLys 404
Db 2089 TTCTAGCAGAGAGAAACAGATGGCAGATATGGCCAAATTAATACTATCTGCTTGAAATTC 2148
QY 404 sThrGluGlnGluGlnGlyGlyGlyGluAlaGlu-----GlyGlyValValVa 421
Db 2149 TGAAGTGAACATAATCGTTGGACTCTCGAGATGAAGAACGATGCTGGCAGATTGATGT 2208
QY 421 iGluGlyThrGlyGluSerLeuProGluLysLeu-----Al 434
Db 2209 GAGTTTCCCAACAGCATATTAAGAGATAGCTTTCGAAAAAGGCTGTACAGATCT 2268
QY 434 aGluProGlnGluValProGlnGluAlaGluProAlaGluGlu----- 448
Db 2269 GGAACCTGGTGAAGTCCCTTGTGATTCTGACGAAGATGGTGAACACAAATCCCACTCAC 2328
QY 449 -----LeuMetLysSerArgLys 455
Db 2329 CAGAGCTCTGCATTATGAAGTTCTGATGTCTTTTATTTAGAGGACAGAGAAAG 2388
QY 455 tCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGlu----- 471
Db 2389 CAAGCTACGTGAGCAGATGAAGAACTCTCTAGTCTTTAGAAAGAACAAATTTTACTC 2448
QY 472 -----GluLysThrLeuProLysHisProGluGlyIleValSerGluValGlu 488

```

```

Db 2449 TTTTGATTGGATAGACATCAACAGACACGACCTAAAGCTTTGCTTGAAGAGTAAATC 2508
QY 488 tLeuSerSer----- 491
Db 2509 CCTCTCTTCACTCTGTGAAGAAATTGGCTTTTCTTGATTGGAGCTCCGATTTGCAAA 2568
QY 492 -----GlnGluArgIleLysValGlnGlySerPro----- 501
Db 2569 TTTTGAACAAACAAAGATTAAGAAAGTTGACTCTCTCCACAGACCTATTTCATCTCTG 2628
QY 502 -----LeuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysG 519
Db 2629 GTACATGAAAAAGAAAGAAATTAGACTGATTCAGAGGAAAGAAATGATGATTAAGAAAG 2688
QY 519 nLysGlyLysArgGlyGlyGlyAspGluGluProGlyGlyLysIleHisThr 539
Db 2689 GGACCATTA-----GAAAGAGAGCAGAGAGAGGAGGCAATGTTGCTTCTC 2733
QY 539 rGluSerProGluSerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGlu 559
Db 2734 TCGTTTTTTACACAGCTCAATCTTTGAACAAGATTCAGAGGATTCAGCATTCAGAGAG 2793
QY 559 uProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAl 579
Db 2794 AAAAGAGAAAGATTCGACTTCTGTTGATGATCTATGGAGACAGACATCTGAGGG 2853
QY 579 aGluGluGlyThrThrSerAspGlyGluLysArgGluGlyIleThrProTTrpAlaSe 599
Db 2854 AGCAAAACAGCACACT-----GATTCATTCAGAACACAGATGTTCTGTTCCATAG 2904
QY 599 r---PheLysLysMetVal-----ThiProLysLysArgValaLysArgProSe 614
Db 2905 CAGATTTATGGAGCTCACACGAGTCAACAGAAAGAAAGAAAGAAAGAACCCGAAACCCAA 2964
QY 614 rGluSerAspLysGluGluGluLeuGluLysValLysSerAlaThrLeuSerSerThrAs 634
Db 2965 AGAGTTTGAAGAAACAGAAAGATACAGAGATTCATCCCAAGCCCAAGATGCTGCTCGA 3024
QY 634 pSerThrValSerGluMetGln-----AspGluValLysThrVa 647
Db 3025 GAATTAAGATTCAGAACTGAAAACTCCACCTTCCGTGGGCTCCCAAGTGCACAGTGT 3084
QY 647 iGlyGluGluGluLysProGluGluProLysArgArgValAspThrSerValSerTrpG 667
Db 3085 AACTCTAGATACGCCCACTACAGACTAGAGAGACCACTGGTGCAAAAACGTAAGGC 3144
QY 667 uAlaLeuIleCysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspAs 687
Db 3145 GCCTTTGTGA-----ACGAGAGAGAGAGCTGGAGCCAGCTTACCTCTCAGAAAG 3195
QY 687 p-----GluGlyGlyProArgThrLeuGlyGlyAspSerHisArgAlaG 702
Db 3196 AGCAAAAGCTGCATCTGAACCTGCTGCCCCCTGT-----GAACACTGGA 3243
QY 702 uGlu-----AlaSerLysAspLysGluAlaGlyThrAspAlaValPr 716
Db 3244 ACAAGTAAGCTGCCCCCAGAGAGAGACCCCGATTAAGAGAGTGGC-----ATGATGCC 3297
QY 716 oAlaSerThrGlnGluGlnAspGlnAlaGlnGlySerSerSerProGluPro----- 733
Db 3298 TCCGGGTGTGGAGAA-----GTTCACTAGGTGACCAAGCCCTTATCT 3342
QY 734 ---AlaGlySerProSerGluGlyGluGlyValSerThrTrpGlu----- 747
Db 3343 GGATGCAAGCCTCCACACTCCGGGGCTCGTTTCCAGGAGAGAGACAGTATGATCC 3402
QY 748 -----SerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuG 764
Db 3403 AGAGCTGACATACCCAGCCACTTTCAAAACCACTCAGAAAGTCTGAGAGAACCAATGA 3462
QY 764 uGluLysAlaGlu-----AspSerSerValGlu----- 773
Db 3463 GCCAAAGGCCGAAAGCCAGAGCCCACTGCAAGTCTGAGCTGATGCAAAACAGAAAGC 3522

```

Qy	1087	ySerGluThrGly-----GlnAlaThrProGluSerLeuGluValProGluValThrAl	1105
Db	4487	---GAATCTGGGGTGGTGCGACGTCTCCCTCGAAAAAGTGAGAGTCCCCAAAGGAGGA	4542
Qy	1105	aAspValAspHisValAlaAlaThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnAl	1125
Db	4543	TGGTTTATCA-----TCCAGTGTGAAAAGTGATGCC	4572
Qy	1125	aValaProGluSerSerGluThrLeuThrAspSerGluThrAsngIysrThrProLe	1145
Db	4573	AGTTGATCCAGACAAGGAACACAGAGAAAGACGCTGTCTGCTCTGGGCGCTCCCCAGA	4632
Qy	1145	uAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSer-----	1162
Db	4633	AGCC---ACCAGTTAGCCCAAGCAGATGAGCTGGAGCAGCGCTGGGAACACATCGCAAA	4699
Qy	1163	---GlnAspSerLysAlaThrAlaAlaValArgGlnSer-----	1174
Db	4690	GCTCGCTGAGCCTCTGCTCTCTGCTCTATTAAGGCAGATGCACAGAGGGCCTTGCCCC	4749
Qy	1175	-----GlnValThrGluGluGluAlaAlaThrAlaGlnLys	1186
Db	4750	AGAGCAGGGAACAGCCTGCACACCAAGCAAGTGAACAGAGCTGGCTGGCGGCATCGG	4809
Qy	1186	s-----GluGluProSerThrLeuProAsnAsnValProAlaGlnGluGluHis-	1202
Db	4810	CTCATCATCAATGACATTTTCTGGGAGCGACAAAATCTCCAGCACCTCCACCTTATCC	4869
Qy	1203	-GlyGluGlu-----ProGlyArgAspValLeuGluProThrGln	1215
Db	4870	TGGAGATCCACAGACATCTGCACCCCGCAGGTGCACAGGCGCTGCAGCCTTCTGA	4929
Qy	1215	ngInGluLeu---ThrAlaAlaValProValLeuAlaLysThrGluValGlyGlnGln	1234
Db	4930	GGAAGGAATGGACAGACATGAGCGCTGTATCTGCATCTCTGGAACTGAGGCTGTACAGA	4989
Qy	1234	uGlyGluValAspTrpLeuAspGlyGluLysValLysGluGlnGluGlnAlaPheValHi	1254
Db	4990	ATCTTCTAGGCCTCCAGTCAATGCTCTGCACCCCTCA-----	5026
Qy	1254	sSerGlyProAsnSerGlnLysAlaAlaAspValThrTyAspSerGluValMetGlyVa	1274
Db	5027	-GCCGGGCCAACAGATACCAAGGAAGCCACAGAGGAATAGCAGTGAACCTCACACTCAGT	5085
Qy	1274	lAlaGlyCysGlnGluLysGluSerThrGluValGlnSerLeuSerLeuGluGluGlyGln	1294
Db	5086	CCAGAACCCAAAGGGTCTAAAGAAAGTGAAGTCACTCTGTTCGGAAGACAAGGGCG	5145
Qy	1294	uMetGluThrAspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGln	1314
Db	5146	CCAGAAACAAACCGCATCCAGCGCCAGCGAACAAACAAGAAAGTGGTGCT----	5200
Qy	1314	uGlyGluGlnGluThrAlaAlaProGluHisGluGlyThrTyGlyLysProValLeuTh	1334
Db	5201	---CCTGTAGAGAGCGATGTCCTCGAATCCAACCAAGCTCAAGT-----	5242
Qy	1334	rLeuAspMetProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSe	1354
Db	5243	---GAGAGTCTCTGCTGCAATAGGGGGACAACAGTA-----CAGCACCCCGA	5286
Qy	1354	rLeuProAspGlnAspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrTh	1374
Db	5287	AGCCCCACAGAGAAAG-----CAGAGTGAGAAACCCCATTCAC	5328
Qy	1374	rValThrGlnThrAla-----GluAlaValGluLysValIleGln	1387
Db	5329	TCTCTCCTCAGTCAATGATCTTCTGACCTAAGCAAGATTCCTCTCCACAGAGAATTCGTCCCA	5388
Qy	1387	uThrValValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeu-----	1405
Db	5389	AGAAATCAGTGTGAGGAAGAGACTCCAACAAAGACTCTGTGCCCCCGACACCTTCCCCC	5448
Qy	1406	-----LeuProAlaGluLysSerAlaThrGlyGlyHisIleTrpThrLeuGlnHisAl	1423


```

Db 5449 ACCTCCCAAGCAGCAGCCGTGATGAGAGCCTCAAGCCAGGTTAGGGTGCATTCCAT 5508
Qy 1423 AGTAspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleValTh 1443
Db 5509 CATTAAGTATGACCCGGTATACCCACCCAGCGAT---CCAGATCTCCATCCACACT 5565
Qy 1443 rProAla-----ProGluSe 1448
Db 5566 GCGTCTGTACTGACAGCAAGAGCTCTACCTCTGCTGCGCTTGCGGGAGATCCACACCA 5625
Qy 1448 rThreuhisProaspLeuGlnGlyIleSerAlaSerGlnArgGluArgGluI 1468
Db 5626 GAGCCCCCTACTAAGAGTACAGAGTGCATCCAGAGGAGAGGCCAGGAGCTAGTC 5685
Qy 1468 uGluAspLysProAspAlaGlyProAspAlaAspGlySer-----SerTh 1484
Db 5686 TACTCCATCTCCAGCTCTTCCCCAGACACAAAGCCTCTGATGTTGACACACAGCTCCAG 5745
Qy 1484 rAlaIleGluLysValLeuLysAlaGluProGluIleLeuGluLeuGluSerLysSerAs 1504
Db 5746 CACCTGAGGAAGATTCTC---ATGAGACCCCAAGTATGTGTCTGCCACCAAGTGTCACTTC 5802
Qy 1504 nLysIleValLeuAsnValIleGlnThrAlaValAspGlnPheAlaArgThrGluThrAl 1524
Db 5803 CACAAGTGTACACACAGCCATTCAGAGCCTGTC-----AGTCTGC 5844
Qy 1524 aProGluThrHisAlaIleAspSerGlnThrGlnValProAlaCysArgLeuAspSerAr 1544
Db 5845 CCCTTGCCACT-----GAGGCCCGGCCCGCCAGTTGACTCTAA 5886
Qy 1544 gGluProAsnArgCysTrpThrLysMetLysAspAlaValMetLysHisProValProG 1564
Db 5887 AAAGCCT-----TTAGAGAGAAAAAAGACAGCCTCCAGTGACAA 5925
Qy 1564 nProArgGluAspLeuGlnValLeuThrValLeuGluAlaTrpAlaGlnProArgLysCy 1584
Db 5926 C---AACTCTGAGATTCACAGCCTCGAGGTGCTGTAGCT---GCTGACAGAGAAAGGT 5979
Qy 1584 sLeuProArgLeuGlnLeuLys-----AlaProValSer 1595
Db 5980 GGCTCCAGTCATTGCTCCCAAAATTACTCTGTATTAGCCGAGATCCTGTACG 6034

RESULT 9
US-09-917-800A-1505
; Sequence 1505, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Caselle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884

```

```

; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1505
; LENGTH: 15231
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 X59601
US-09-917-800A-1505

Alignment Scores:
Pred. No.: 1,63e-20 Length: 15231
Score: 427.00 Matches: 366
Percent Similarity: 33.95% Conservative: 258
Best Local Similarity: 19.91% Mismatches: 745
Query Match: 5.29% Indels: 471
Db: 10 Gaps: 70

US-09-902-432-4 (1-1596) x US-09-917-800A-1505 (1-15231)
Qy 3 AlaGlySerSerThrGluGlnArgSerProGluGlnProAlaGlySerAspThrProSer 22
Db 4467 GCTGGCTAGACACAGCGGCGCA-----GAGCCGAGAGCCCTTGCC 4508
Qy 23 GluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSerGlyAlaGlyAspPro 42
Db 4509 CGA-----GTTGAGAGCGCGCGCTGGAGAAAGCAGCGGCGCTGGTGAAGCCCA 4556
Qy 43 -AlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAs 62
Db 4557 TGCCAGAGCGCAAGGACAGAGCGGCGTGAAGCAGCAAGTGCAGCGCGCATCAGGA 4616
Qy 62 nGlyValAla-----GluGlnGlyAspValHisValGlnGluGluAsnGlnGlyG 80
Db 4617 GAGGTGAGCGCGCGCGAGAGGCGCGGTGAGCCACAGCAAGAGCGCAGCATCA 4676
Qy 80 nGluGlu-----GluValValAs 86
Db 4677 AGAGAGGCTGCAGCATTCGCGCAAGCTCAGAGGACAGATCCAGGCCAAGGCCACGA 4736
Qy 86 pGluAspValGlyGlnArgGluSerGluAspValArgGluLysAspArgValGluGluMe 106
Db 4737 GGTGAGGCTGCAGAGCCAGCGCATCGCATTCAGAGAGAGATCCCGTAGTCCGCT 4796
Qy 106 ValAlaAsnSerThrAlaValAlaGluAspIleThrLysAspGlyGlnGluThrSerG 126
Db 4797 GCAGCTAGAGACT-----GAGCGTCAGCGTGAAGGCGCGAGATGAGCTGCA 4847
Qy 126 uIleIle-----GluGlnIleProAlaSerGluAsnAsnValGluGluMetVal 142
Db 4848 GCGCTCGGTGTCAGCGGCTGAGAGGACAGAAAGCAGAGCGCGCGATGAGAGAAC 4907
Qy 142 LginProAlaGluSerGlnAlaAsnAspValGlyPheLysValAlaPheLysPheValG 162
Db 4908 CGAGGCTTCGAGAGCGAGTGCAGATGAGAGCCCAAGCAAAAGCGAGCGGAGCCGA 4967
Qy 162 yPheLysPheThrValLysLysAsp-----LysAsnGlnLysSerAspThrValG 179
Db 4968 GCTGGCCCTGCTGTGAAGGACAGAGCGGAGGACCGGAGAGAGCGGAGCCCTGCA 5027
Qy 179 nLeuLeuThrValLysLysAspGluGlyGluGlyValGlu-----AlaSe 194
Db 5028 GCGCTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5087
Qy 194 rValGlyAlaGlyAspHisGlnGlnProSerValGluThrAlaValGly----- 210
Db 5088 GGCAGAGAGGCTCCCAAGTGCAGTAGTACCTGTGAGACAGCGCAGCTAGTCAGAAAGT 5147
Qy 211 -GluSerAlaSerLysGluSerGluLeuLysGlnSerThrGluLysGlnGluGlyThre 230

```

Db	5148	CGAGCTGGAGCAGCAGCGCTCCGCTCTTTCGACAGAGAAGACCGCACAGTCTGGAGCGGACGCCT	5200
Qy	230	uLyEglnGluN---SerSerThrGluileProLeuGlnAlaGlu-----SerAs	246
Db	5208	GCAGGAAGACACGTGACAGTACACAGCTGCGGGAGGAGCGGAACGGCGGGCACAGCA	5267
Qy	246	pGlnAlaAlaGluGluAlaLysAspGluGlyGluGluLysGlnGluLysGluProTh	266
Db	5268	GCAGGCTGAAGCCGACGAGCCCGTGTAGGAAGCCGAGCGGAGCTGGAGCGCTGGCAGCT	5327
Qy	266	rLySerProGluSerProSerSerProValAenSerGluThrThrSerSerPheLyLy	286
Db	5328	GAAGCCAAATAGCGCTCGCGCTCGAGCAGAGGAGTGGCAGACAGAGAAG	5387
Qy	286	sPhePheThrHisGlyTrpAlaGlyTrpArgLyLysThrSerPheLyLysSerLySgI	306
Db	5388	CCTG-----GCCAGGCGCATGCGGAGAGCAGACGGA	5420
Qy	306	uAspAspLeuGluThrAlaGluLysArgLySglnGlnGluAlaGluLysValAspGluG	326
Db	5421	AGAGCGAAGACGGGAAGCCCGCGCGGCGCAAGCAGAGGACGAGCGCTGCGGCACGC	5480
Qy	326	uGluLySglnLysThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspG	346
Db	5481	AGAGCTGGTGAGCAGAG---CTGGAGAAGCAGCGGCGAGCTGACAGCGGCGCACCGGCCA	5537
Qy	346	nAlaArgLeuSerAlaAspTyrgLyuLysValGluLeuProLeuGluAspGlnValGlyAs	366
Db	5538	GCAGCGCTGGCTGCCAGCAGAGCTGATTGCTTGGCGGCAGAGCAGGACAGAGTGTA	5597
Qy	366	pLeuGluAlaSer---SerGluGluLysCysAlaProLeu-----AlaThrGluVa	382
Db	5598	GCATCAGCGGCAGCTGTGTGAGGAAGAGCTGCGCGGCTACAGCACGAAGCGACAGCAGC	5657
Qy	382	lPheAspGluLysMetGluAlaHisGlnGlu-----ValValAlaGluValHisVa	399
Db	5658	CACACAGAGCGGCAGAGCTGGAGCTGAGCTGGCAGAGTTCGGCAGAGATGGAGGT	5717
Qy	399	lSerThrValGluLysThrGluGluGlnGlyGlyGlyGluAlaGluGlyGlyVa	419
Db	5718	ACTGCTGCCAGCAGGACGACGCGCCGAGAGAGTCTCGCTCCACCATGGA-----	5769
Qy	419	lValValGluGlyThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGluVa	439
Db	5770	-----AAGTCCAAGCAGAGGCT	5786
Qy	439	lProGlnGluAla-----GluProAlaGluGluLeuMetLySerArgGluMe	455
Db	5787	GGAACTGGAGGAGGCGGCTTCAGAGAGCTGTCTAGGAGAGCTGGCTGCATCAGTAGGC	5846
Qy	455	tCysValSerGlyGlyAspHisThrGlnLeuThrAspLeuSerProGluGluLysThrLe	475
Db	5847	GGCCGAGGAGCAGGCGGCACCGGAGTGTGCC-----GAGGAGGACGCGGC	5894
Qy	475	uProLyHisProGlu-----GlylleValSerGlu---ValGluMetLeuSerSerG	492
Db	5895	ACGCCACGCGGCGGACGAGCGAGTGTCTACGAGAGAGCTGGCTGCATCAGTAGGC	5954
Qy	492	nGluArgIleLys-----ValGlnGlySerPr	501
Db	5955	CACAAGCTCTCAAGACGGAGCGACAGATTGCATCTAAAGAGGAGGCGCGACAGCAGCG	6014
Qy	501	oLeuLyLysLeuPheSerSerSerGlyLeuLyLysLeuSerGlyLyLysGlnLySgI	521
Db	6015	CCTGAGCGGCTGGCTGAAGATGAGCGCTTCAGCGGCGCGGCTGGGAGGAGCAGGCGAGC	6074
Qy	521	yLyAsrGlyGlyGlyAspGluGluProGlyGluTyrglnHisIleHisThrGluSe	541
Db	6075	ACAGCACAGGCGACATAGAGGAGCGCTCGGCCAG-----CTGCCAAGGCATC	6125
Qy	541	rProGluSerAlaAspGluGlnLySgLy-----GluSerSerAlaSerSerProGluG	559
Db	6126	CGAGAGCGAGCTGGAGCGACAGAAGGGTGTGTGGAGATATCCCTCGGCGGACGGCGGCA	6185

```

Oy 869 IHisThrValSerValAlaValIleAspGlyThrArgAlaValIHisSerValGluGluThr 889
Db 7147 -----CTCAAGGCTCGCATTGAAGCTGAAACCG 7175
Oy 889 GSerProSerTrpIleSerAlaSerValThrGluPro---LeuGluHisThrAlaGlyI 908
Db 7176 GGCATCTACTCTTCCTGTCAGACAGACACACACAGCCCTTCCTGGAGAGAGAGCCGAGAA 7235
Oy 908 uAlaMetProProValGluGluValThrGluLysAspIleIleAlaGluGluThrProVa 928
Db 7236 GATGAAACAGGTGGCGACAGAGAGAGCTGCACGCTTGAGCGTGAAGCTCCAGAGAGCAGCAAG 7295
Oy 928 lLeuThrGln-----ThrLeuProGluGluLysAspAlaHisAspMetVal 944
Db 7296 GCTGGCGAGCTACGCGCAGAGAGAGCTGGCCAGAGAGCGGCGCTGGCGAGAAAGATGCT 7355
Oy 944 lHisSerGluValAspPheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLe 964
Db 7356 GAAGAGAAAGATCGACGCGCTGCAGAGAGCCACAAAGCTCAAGCTCAAGCTGAGCTGCT 7415
Oy 964 uArgThrGluGluValThrGluAlaSerGlyAlaGluGluThrThrAspMetValSerAl 984
Db 7416 GCAGCAGCAGAGAGAGCTG-----GCACAGCAGCAGCGCCGCGCTGCAGAGC 7463
Oy 984 aValSerGluLeuThrAspSerProAspThrThrGluGluAlaThrProValGluGluVa 1004
Db 7464 G-----GACAAAGAGCAAAATGCTCAGCAGCTGAGTGAAGGA 7499
Oy 1004 lGluSerGlyValIleAspThrThrGluGluGluGluArgGluThrGluAlaIleLeuGlnAl 1024
Db 7500 GACACAGGCTTCCACAGCGGACCTGTGAGGCTAGAGCGGCGGCACTGAAATGAGCC 7559
Oy 1024 aValAlaAspLysValLys-----ValProAlaThrGluThr-----ValGlnArgThrGl 1045
Db 7560 AGAGCTGAACCGCTCAAGATTGCGCATGTGATGATGACCGGAGCTCAGGCGCTGCAGCA 7619
Oy 1031 -GluGluSerGln-----ValLeuAlaSerGluLysGluLysAspValMe 1071
Db 7620 GGAGGATGCCAGCGCTTCCGAGAGAGCTGAGAGATCGCGAAAGCTGCACCGCAC 7679
Oy 1035 -----ValProAlaThrGluThr-----ValGlnArgThrGl 1045
Db 7680 TGAATCGCTACACAGAGAGAGAGATGACATTGGTCAGACTTCGAGATCGACGACAGCA 7739
Oy 1045 ySer-----LysAlaLeuGluLysValGluGluValGluGluAspSerGl 1060
Db 7740 GAGTGACCAAGATGCCAGCGCTGTGAGGAGCGCATGCTGAGCTGAGACGCGAGAGA 7799
Oy 1060 u-----ValLeuAlaSerGluLysGluLysAspValMe 1071
Db 7800 GAAGCTCAAGCAGAGCGAGAGATTACTGCAGCTCAAGCTCAGAGATCGACACTGTGCA 7859
Oy 1071 tProLysGluProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGl 1091
Db 7860 GCAGAGCAGATATCTGCAGAGAGACACAGGCGCTGCAGAAAGAGCTTCTCTGAGAGAGA 7919
Oy 1091 yGlnAlaThrProGluSerLeuGluValProGluValThrAlaAspValAspHisValAl 1111
Db 7920 CAGCTTGCTGCACACCGGAAGCTTCATCGACGAGGAAAGGCCAAGCTGAGAGAGCTTTT 7979
Oy 1111 aThrCysGlnValIleLysLeuGlnGlnLeuMetGlu-----GlnAl 1125
Db 7980 CCAGAGCAGGTGGCAAAAGCAAAACAGCTGCAGAGGAGGAGCAGCAGCGCAGCAGAGCA 8039
Oy 1125 aValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThrProle 1145
Db 8040 GATGGAGCAGGAAAGACAGAGAGCTGTGGCCAGCATGGAG-----GAGGCCGCGAG 8090
Oy 1145 uAlaAspSerAspThrAlaAspGlyThr-----1154
Db 8091 GGGGCAAGCTGAGGAGAGAGAGGTGTGAGCGCCAAAGCAAGAGAACTGACAGGCTGGA 8150
Oy 1155 -----GlnGlnAspGluThrIleAspSerGlnAspSerLysAlaThrAlaAlaVa 1171

```

```

Db 8151 GCAGACGGCGGACAGCAGAGGAAACTACTGGCAGAGAGAGAACAGAGGCTCGCGGAGCG 8210
Oy 1171 lArgIleSerGlnValThrGluGluGluAlaIleThrAlaGlnLysGluGluProSerTh 1191
Db 8211 GCTCAGAGCGCTTGAGAGAAAGACACCGAGCTGCTGGCCGCACTGAGAGATGCGCAC 8270
Oy 1191 t-----LeuProAsnAsnValProAlaGlnGluGluHisGl 1203
Db 8271 CTCCAGGCTGCTGCCAAGAAAGCATGCCCAAT-----8304
Oy 1203 yGluGluProGluArgAspValIleGluProThrGlnGlnGluLeuThrAlaAlaAlaVa 1223
Db 8305 -----GCGCGCAGCAGCATTTGATGCGCCCTCCATGGAG-----8337
Oy 1223 lProValLeuAlaLysThrGluValAlaGlyGlnGluGlyGluValAspTrpLeuAspGlyI 1243
Db 8338 -----GCCAGCGCGAGTACACTTTGAGGAGATTACGTGAGAGAGTGCAGAGCTCA 8387
Oy 1243 uLysValLysGlu-----GluGlnGluValPheValHisSerGlyProAsnSe 1259
Db 8388 GCAGCTACAGAAACAGGCAATTCTGAGCATGAGAGAACTGCAGCGCTTGCACAGAGGTCA 8447
Oy 1259 rGlnLysAlaAlaAspValThrTyraAspSerGluVal-----Me 1272
Db 8448 CACCAAGGTGCTGAGCTGCTACAGCAGCGGAAAGATGTGCGCCACTGAAGGCGGCGAG 8507
Oy 1272 tGlyValAlaGlyCysGlnGluLysGluSerThrGluValGlnSer-----1287
Db 8508 CAGATCGCAGAGATTGCTCTGMAAGCCACCAATGAGAACTGATGTCACAGAGCTT 8567
Oy 1288 -----LeuSerLeuGluGluGluGluMetGluTh 1297
Db 8568 ACAGCGCAGCTGTCAAGCCCTGGAACAGCCCTTATCTTACTTGAAGGCCAGAGCTTC 8627
Oy 1297 t-----AspValGluLysGluLysArgGluThrLysProGluGluLysSerGl 1313
Db 8628 GGGCTTCTGCTGACCTCTGTCGGAACCGGCGCTGAGGCTCAATGAGCTGTGAAGA 8687
Oy 1313 u-----1313
Db 8688 GGGTGTGGTGGTCCCGAGCTGCACACACAGCTGCTGTCACTGAGCGTGCATCAG 8747
Oy 1314 -----GluGlyGluGlnGluThrAlaAlaProGluHisGlu-----1325
Db 8748 CTACAAAGACCTTACACAGAGAGAAACAGAT--CTCTCTTCCAGGCGCATGAAAGAGACC 8806
Oy 1326 -----GlyThrTyrgLysProValLeuThrLeuAsp-----1336
Db 8807 TCATTGTCAAGGACCATGAGATCCGCGCTGTGGAAGGCCAGATGCGCACAGGTGGCATCA 8866
Oy 1337 -----MetProSerSerGluArgGlyLysAlaLeuGlySerLeu-----1349
Db 8867 TTGACCTGTACACAGCAGCAGCGTGTCCGTTGAGACGTGCTTACAGGTGGTACTTCTG 8926
Oy 1350 ---GlyGlySerProSer-Leu-----ProAspGlnAspLysAlaGlyCysIleGluV 1366
Db 8927 ATGAGAGATGAACCTGTGCTGCTGCTGACCCAAAGGAGTACACCAAGGCGCTTTTGAAC 8986
Oy 1366 aGlnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValI 1386
Db 8987 CCAACACTCAC-----GAGAACTCTCA 9007
Oy 1386 lGluThrValValIleSerGluThrArgLysGluSerProGluCysValGlyAlaHisLeuL 1406
Db 9008 CGTACTCGACGTGCTGAGAGCGGTGTGTGAGAGACCCGAG--ACAGGCTGCGGCTTCC 9064
Oy 1406 euProAlaGluLysSerSerAlaThrGlyGlyHis-Tyr---ThrLeuGlnHisAlaGlu 1424
Db 9065 TGCCACTCAGACAGCAAGAGCTCCAAAGGGTGTGAGCTGTGTACTGACACTGACAGAGAGCC 9124
Oy 1425 AspThrValProLeuGlyProGluSerGlnAlaGluSerIleProIleIleValThrPro 1444

```

Db 9125 GTGACGCTCTCGAAGGCCACAGTGTCTGCACCATTCGCCAAGTTCACGGGCAAGACCG 9184
Qy 1445 AlaProGluSerThrLeuHisProAspLeuGlnGlyGluLeuSerAlaSerGlnArgGlu 1464
Db 9185 TGACCATCTGGGAGA-----TCATCAACTCAGAGTACT 9217
Qy 1465 ArgSerGluGluAspLysProAspAlaGlyProAspAlaAspGlyLysGluSerThr 1484
Db 9218 TCACAGCGGAGCAGCAGCGGAGCTGCTCCGCGAGTTCGCCAGC-GGCCGC-----ATC 9270
Qy 1485 AlaileGluLysValLeuLysAlaGluProGluLeuLeuGluSerLysSerAsn 1504
Db 9271 ACGGTGGAGAGATCATCAGATTGTCATCAGCGTGGTAGGAAACACAGCGGAGGGC 9330
Qy 1505 Lysile-ValLeuAsnValileGlnThrAlaValAspGlnPheAlaArgThrGluThrAl 1524
Db 9331 CAGCTCTGCTTGT----- 9343
Qy 1524 aProGluThrHisAlaTyRaspSerGlnThrGlnValProAlaCysArgLeuAspSerAr 1544
Db 9344 -----AGGGCCTCCGTGCCCTTGTGCTGTGCAGAGC-- 9376
Qy 1544 gGluProAsnArgCysTrpThrLysMetLysAspAlaLysMetLys 1559
Db 9377 -----TGCTGGACA---GTGGAGTCATCATGTATGAAG 9406

RESULT 10
US-09-954-456-1601
; Sequence 1601, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1601
; LENGTH: 14800
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1601

Alignment Scores:
Pred. No.: 6,73e-20 Length: 14800
Score: 418.00 Matches: 334
Percent Similarity: 35.65% Conservative: 255
Best Local Similarity: 20.22% Mismatches: 701
Query Match: 5.18% Indels: 364
DB: 10 Gaps: 61

US-09-902-432-4 (1-1596) x US-09-954-456-1601 (1-14800)
Qy 5 SerSerThrGluGlnArgSerProGlu-GlnProAlaGlySerAspThrProSerGluLe 24
Db 4187 AGCAGCAGCGGGCAGAGGAGCGCGAGCGCTGCGCGAGGTGGAGGCGCGCTGAGAGAAGC 4246
Qy 24 uValLeuSerGlyHisGly-----ProAlaAlaGluAlaSerGlyAlaAlaGlyAspPr 42
Db 4247 AGCGGAGCTGCGCGAGGCGCACGCCCAAGGACACAGCGCGGAGCGGAGGCGAAGG 4306
Qy 42 oAlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSerSerValAs 62
Db 4307 AGCTGAGCAGCGCATGCA-----CGAGAGGTGTGTGCGCGGGAGGAGCGG 4354
Qy 62 nGly-ValAlaGluGlnGlyAspValHisValGlnGluAsnGlnGluGlyGlnGluG 82
Db 4355 CGGTGACGCGCAGCAGCAGCAAGCGCAGCATTCAGGAGGAGCTGCAGCAGCTCGCGCAGA 4414
Qy 82 luGluValValAspGluAspValGlyGlnArgGluSerGlu----- 95
Db 4415 GCTCGGAGCGGAGATPCCAGGCCAAGCGCCGCGCAGCAGAGCGCGCTGAGCGCAGCGGC 4474
Qy 96 --AspValArgGluLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluA 115
Db 4475 TGGCATCGAGGAGGAGATCCGGTGTGTGCGCTGCAGTTGAGGAGCCCGAGCGGCAGC 4534
Qy 115 spileThrLysAspGlyGlnGluGluThrSerGluile---lleGluGlnIleProAlas 134
Db 4535 GTGGCGGGCTGAGGGGAGCTGCAGCCTCGCTCACGGCGGAGGAGCTGAGGCAC 4594
Qy 134 erGluAsnAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyP 154
Db 4595 AAAAGCAGCAGCGCAGGAGGAGCGCGCTTGGCGAGGCGAGGTGCAGGACGAGAGCC 4654
Qy 154 helysLysValPheLysPheValGlyPheLysPheThrValLysLysAsp-----L 171
Db 4655 AGCGTAAGCGGCGAGCGGAGGTGGAGCTGCGCTCGCGCTGAAGCGCGCAGCAGCGCG 4714
Qy 171 yeAsnGluLysSerAspThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyA 191
Db 4715 CGCGCGAAGAGCAGCGGGCCCTGCAGGCCCTGGAGGAGCTGCGGCTGCAGCGGAGGAGG 4774
Qy 191 laGlu-----AlaSerValGlyAlaGlyAspHisGlnGluProSerValG 206
Db 4775 CGGAGCGCGCTCGCGCAGCGCGAGGTGGAGCGCGAGCGCGAGGTACAGGTGCGCTGG 4834
Qy 206 luThr-----AlaValGlyLysSerAlaSerLysGluSerGluLeuLysGlnS 222
Db 4835 AGACGCGCAGCGCAGTGCAGAGCGCGAGCTGCAGAGCAACCGCGCTCTCTCGCCGAGA 4894
Qy 222 erThrGluLysGlnGluGlyThrLeuLysGlnGluGln---SerSerThrGluIleProL 241
Db 4895 AGACGCGCAGCTGGAGCGCTCCCTGCAGGAGGAACACGTGCTGTGGCACAGCTGCGGG 4954
Qy 241 euGlnAlaGlu-----SerAspGlnAlaAlaGluGluAlaLysAspGluGlyG 258
Db 4955 AGGAGCTGAGCGCGCGGCGCACAGCAGCGCGAGCGCGCGCGCGCGAGGAGGAGCAG 5014
Qy 258 luGluLysGlnGluLysGluProThrLysSerProGluSerProSerProSerProValAsnS 278
Db 5015 AGCGGAGCTGGAGCGCTGCGCAGCTCAAGGCCAACCGCGCTACGCGCTGCGGCTCGAG 5074
Qy 278 erGluThrThrSerSerPheLysLysPhePheThrThrHisGlyTrpAlaGlyTrpArgLysL 298
Db 5075 CGGAGGAGGTGGCGCAGCAGAGAGCGCTG-----CGCG 5107
Qy 298 ysThrSerPheLysLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluG 318
Db 5108 AGCGCGAGCTGAGAGCAGAGAGGAGGCGGAGCGCGCGCGCGCGCGCGCGCGAGG 5167
Qy 318 lngluAlaGluLysValAspGluGluGluLysGluLysThrGluProAlaSerGluGluG 338
Db 5168 CGGAGGAGCAGCGCTCGCGCAGCGGAGCTGGCTGAACAAGAG---CTGAGAGAGCAGC 5224

[illegible]

QY	630	LeuSerSerThrAspSerThrValSerGluMetGlnAspGluValLysThrValGlyGlu	643
Db	6262	CAGAAAGCGCGCTG-----CAGGGCGAAGAAAGACACAGCGCTTCGCGGTG	6306
QY	650	GluGlnLysProGluGluGluProLysArgValAspThrSerValSerTrpGluAlaLeu	669
Db	6307	CAGCAAGAAAGACGAGAGCTACACAGACAGCGCTGCAGACAGACAGAGC-----CTGCTG	6360
QY	670	IleCysValGlySerSerLysLysArgAlaArgGlyAlaSerSerSerAspAspGluLys	689
Db	6361	GACCAGCTGCGCGCGGCGAGCGGAGCGCGCCCGCGCGCGCTGAGAGAGCGGAGAGGACC	6420
QY	690	-----GlyProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLys	706
Db	6421	CGGGTCAGCGGAGCGCTGAGGGCGGCGAGGCCCGCGCGAGTGAAGAGAGCGCGAGCGG	6480
QY	707	AspLysGluAlaGlyThrAspAlaValProAlaSerThrGluGluGlnAspGlnAlaGln	726
Db	6461	CTGAAGCAGCTCG-----GCAAGAGACAGAGCGACAGAGCCCGGCGCTGAG	6522
QY	727	GlySerSerSerProGluProAlaGlySerProSerGluGlyValSerThrTrp	746
Db	6523	GCACAGCGCGGTGCAGAG-----	6540
QY	747	GluSerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluLys	766
Db	6541	-----AAGCTCCGCAAGAGCGCCGACAGAGAGCGC	6570
QY	767	AlaGluAspSerSerValGluGlnLeuSerThrGlnIleGluProSerArgGluGluSer	786
Db	6571	CGCGCGCGGGACAGCGCGAGAGCGCGCGCTTCGCGACAGACAGACAGCTGACCGCGAG	6630
QY	787	TrpValSerIleLysLysPheIleProGlyArgArgLysLysArgAlaAspGlyLysGln	806
Db	6631	ATGGAAGAAGCATTAAGAAATTGGCGGACAGACAGCTGCGGCAAGAGCGGAGTGGACGAG	6690
QY	807	GluGlnAlaThrValGluAspSerGlyProValGluIleAsnGluAspAspProAsnVal	826
Db	6691	GAGCTGACCAACACTGCGCG-----CTGCAGCTGGAAGACACCGAC-----	6729
QY	827	ProAlaValValProLeuSerGluTrpAlaAlaValGluArgGlu-----Lys	842
Db	6730	-----CACCGAAAGACTGCTGTGAAGAGAGACTGCAGCGGCTGAAG	6771
QY	843	MetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyValAlaVal	858
Db	6772	GCGAGGCGCCAGCGAGCGCGGACCGGACCGGAGCGAGCGTGAAGAGAGACTTCTTCGGTG	6831
QY	859	TyrValSer--GluGluLeuSerLysThrLeuValHisThrValSerValAlaValIle	877
Db	6832	CGCGTGCAGATGATGAGGAGCTGAGCAAG-----	6858
QY	878	AspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIleSerLase	897
Db	6859	-----CTCAAGGGACCGCATCGAGGCTGGAACCGCGCATCATTTGCGTGCAGAGGAC	6912
QY	898	ValThrGluProLeuGluHisThrAlaGlyAlaIleMetProProVal---GluGluVal	916
Db	6913	AATACGACGCGCTTCCTCGCAGAGAGAGCGCTGAGAGATGAACAGAGTGGCGGAGGAGCC	6972
QY	917	ThrGluLysAspIleIleAlaGluGluThrProValLeuThrGln-----Thr	932
Db	6973	GCGCGCGCTGAGTGTGGCGGCCCAAGAGCGTGGCGCATCGGCGGTGGCAGAGAGGAC	7032
QY	933	LeuProGluGlyLysAspAlaHisAspAspMetValThrSerGluValAlaAspPheThrSer	952
Db	7033	CTGGACACGACGCGCGGCTTCGCGAGAGAAATGCTTCAGAGAAAGATGACGCGGTGCAG	7092
QY	953	GluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAla	972
Db	7093	GAGGCGACGCACTCAAGGCTGAGGCGGAATCTGCTCAGACAGCAAGAGAGCTTCGCGAG	7156
QY	973	SerGlyAlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerPro	992

Db 7153 GAGCAGCGCGCGCGCTGCAGGAG----- 7176
Qy 993 AspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGlu 1012
Db 7177 GACAGGAGCAGATGGCGCAGCAGCTGGCGAGGAGAGCAGCGGCTTCCAGCGGAGCGCTG 7236
Qy 1013 GluGluGluArgGlnThrGlnAlaIleGlnAlaValAlaAspLeuValys----- 1030
Db 7237 GAGCGCAGCGCGCAGCGCAGCTGGAGATGAGCTGAGCGCTGAGCGCCTCAAGCTCGT 7296
Qy 1031 -----GluGluSerGlnValProAlaThrGlnThrValGlnArgThrGlySer 1046
Db 7297 GTGCCGAGATGACCCAGCGCCAGCGCGCTGAGGAGAGCGCCAGCGCTTCCGGAAG 7356
Qy 1047 LysAlaLeuGluValGluGluValGluGluAspSerGluValLeuAlaSerGluLys 1066
Db 7357 CAGCGGAGGAGATCGGTGAGAGCTGCACCGCAGCGAG-----CTCGCCACCCAGGAG 7410
Qy 1067 GluLysAspValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGln 1086
Db 7411 AAGGTGACCTGTGTGCAGACACTGGAGATCCAGCGCAGCAGAGTGACCATGTGCCGAG 7470
Qy 1087 GlySerGluThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrAlaAsp 1106
Db 7471 -----CGCCTGCGGGAGCGCATCGCTGAG 7494
Qy 1107 ValAspHis-----ValAlaThrCysGlnValIleLysLeuGlnGlnLeu---MetGlu 1123
Db 7495 CTGGAGCGTGAGAGGAGAGCTCCACAGCAGGAGGCCAACTGCTGCAGCTCAAGTCTGAG 7554
Qy 1124 GlnAlaValAlaProGluSerSerGluThrLeuThrAspSerGluThrAsnGlySerThr 1143
Db 7555 GAGATGCAGCGGTGTCAGGAGCAGCTGCTGCAGGAGCAGCGGCCCTGCAGCAAGC 7614
Qy 1144 ProLeuAlaAspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGln 1163
Db 7615 TTCTCTCTGAAAGGACACCTG-----CTACGCGGGAGCGCTTCATCGAGCAGGAG 7668
Qy 1164 AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGluAlaAlaThr 1183
Db 7669 AAGGCCAAGCTG-----GAGCAGCTCTCCAGGACGAGGTGGCCAG 7710
Qy 1184 AlaGlnLysGluGluProSerThrLeuProAsnValProAlaGlnGluGluHisGly 1203
Db 7711 GCACAGCAG-----CTGCGTGAGGAG----- 7731
Qy 1204 GluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThrAlaAlaVal 1223
Db 7732 CAGACGGCGCAGCAGCAGATGGAGCAGGAGCGGCTGTGGCCAGCATGGAG 7791
Qy 1224 ProValLeuAlaLysThrGluValGlyGlnGluGly-----GluVal 1237
Db 7792 GAGCGCGCGCGCGCAGCATGAGCGCGGAGGCGCTGGCGCAGCAGGAGGAGCTG 7851
Qy 1238 AspTrpLeuAspGlyGluLysValLysGluGlnGluValPheValHisSerGlyPro 1257
Db 7852 CAGCAGCTGGAGCAGCAGCGG-----CGGCAGCAGGAGGAGCTGTGGCTGAGGAGAACCG 7908
Qy 1258 AsnSerGlnLysAlaAlaAspValThrTyrAspSerGluValMetGlyValAlaGlyCys 1277
Db 7909 AGCTGGTGAGCAGCTGACCTGCTGGAGGAGCAGCAGCGCGCGCTGGCGCACTCA 7968
Qy 1278 GlnGluLysGluSerThrGluValGlnSerLeuSer----- 1289
Db 7969 GAGGAGTCACTGCTCGCAGGTGGCTGCCACAAGACCTGCCCAATGCGCGGATGCA 8028
Qy 1290 -----LeuGluGluGlyGluMetGluThrAspValGlnLysGluLysArgGlu 1305
Db 8029 CTTGATGGCCCGCGCGCAGGAGCAGCGGAGCAGCAGCTTCGATGCTGCGCGGAG 8088
Qy 1306 ThrLysProGluGlnValSerGluGluGlyGlnGluThr----- 1319

Db 8089 GTCTCAGCTCAGAGGCTGCAGGAGCGCGCAT--CCTGAGTGGGAGGAGCTGCAGCGGTT 8147
Qy 1320 ---AlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuAspMetPro 1338
Db 8148 GGCAGCGGCGCACACCGGTGA----- 8171
Qy 1339 SerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAspGln 1358
Db 8172 -----CGAGCTCGCACGCGCGGGAAGACGTGCCCACTACCTGCGGCGCGCAG 8222
Qy 1359 AspLysAlaGlyCysIleGluValGlnValGlnSerLeuAspThrValThrGlnThr 1378
Db 8223 TATCGAGGGCTGCTGAAGCCAC-----CAATGAGAA 8258
Qy 1379 AlaGluAlaValGluLysValIleGluThrValValIleSerGluThrGlyGluSerPro 1398
Db 8259 GCTGAGTGTTCAGCGCGCTGCAGAGGAGCTGTGAGTCCCGGCGCAGCCCTCATCCT 8318
Qy 1399 Glu----- 1399
Db 8319 GCTGAGGCGCAGCGCGCTCAGGCTTCTGTGGACCCCTGTGCGGAACCGCGCGCTGAC 8378
Qy 1400 -----CysValGlyAlaHisLeuLeuProAlaGluLysSer 1411
Db 8379 CGTCAACGAGGCTGTGAAGAGGCTGTGTGGCCCCCGAGCTGCACACCAAGTGTCTGC 8438
Qy 1412 SerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrValProLeuGlyPro 1431
Db 8439 GGCAGGCGCGCGCTCACTGGCTACAGGACCCCTACACTGGCCAGCAGATCTCTCTT 8498
Qy 1432 GluSerGlnAlaGluSerIleProIleIleValThrProAlaProGluSerThrLeuHis 1451
Db 8499 CCAAGCATGTCAGAAAGG-----CCTCATGCTCGGCGAGCAGCGCAT 8540
Qy 1452 ProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluArgSerGluGluGluAspLys 1471
Db 8541 CCG----- 8543
Qy 1472 ProAspAlaGlyProAspAlaAspGlyLys 1481
Db 8544 CCGTGTGGAGCGCCAGATCGCCACGCGCGG 8573
RESULT 11
US-09-815-242-4580
; Sequence 4580, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308

Qy	602	g-----MetValThrProLysLysArgValaArgPr	613
Db	3783	AAATGCTAAAAAATGATGTTGATCAAGCTGTGCAACTCTCAAAATCAAGCAATTTGATAATAC	3842
Qy	613	oSerGluSerAspLysGluGlu	620
Db	3843	AACTGGTGCTACAACTGAAGAAAAAATGCAGCAAAAGATTGTTTTAAAAAGCTAAAGA	3902
Qy	621	-----GluLeuGluLysValLysSe	627
Db	3903	AAAAGCGTATCAAGATATCTTAAATGCACAAACAACTAACGATGTTACGCCAAATTAAGA	3962
Qy	627	rAlaThrLeuSer-----SerThrAspSerThrValSerGluMet-----G1	641
Db	3963	TCAAGCAGTGTGCTGATGTTCAAGGTATTACTGCAGATACAAACATTAAGATGTTGCCAA	4022
Qy	641	nAspGluValLysThrValGlyGluGluGlnLys-----	652
Db	4023	AGATGAATTAGCAACAAAGCAAGGAAACAAAAGCACTTATTGCACAACTGCAGATGC	4082
Qy	653	-----ProGluLupProLysArgValAspThrSerValSerTrp-----	666
Db	4083	GACTACTGAAGAAAAGAACACAGCAATCAACAGTAGATGCAGATTTAACGCCAAGTAA	4142
Qy	667	-----GluAlaLeuIleCysValGlySerSerLysLysArgAlaAr	680
Db	4143	TCAAAATATTGAAATATGCACAGTCATCGATGTATAACACTGCACAAAGATAATGCAAT	4202
Qy	680	gLysAla-----SerSerSerAspAspGluGlyGlyProArgThr-----	693
Db	4203	TCAAGCAATTGACCCAATTCAAGCATCAACAGATGTTTAAACGAATGCAGACGCCGAATT	4262
Qy	694	-----LeuGlyGlyAspSerHisArgAlaG1	702
Db	4263	GCTAACTGAAATGCMAAATAAATACTGAAATACTTAATTAATAGACTACTAATGA	4322
Qy	702	uGluAlaSerLysAsp-----LysGluAlaGlyThrAspAlaVa	715
Db	4323	AGAAAAAGGTAAACGATATTGGACCAGTTAGAGCTGCATATGAAGAAGGTTTAAATATAT	4382
Qy	715	lProAlaSerThr-----GlnGluG1	722
Db	4383	TAATGCGACCAACTACTACAGTGATTAACACTGCTAAAGATACAGCAGTACAAAAAGT	4442
Qy	722	nAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGlyGluG1	742
Db	4443	TCAACAACCTTCATGCAAACTCCTTTAAGAAGCCAGCAGGTAAACTGCTTTAGATCAAGC	4502
Qy	742	yValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLysSerLy	762
Db	4503	TGCAGCTGAT---NAGMAAACAAATAGACAAACACCAATGCGTCAACAAAGAANAAT	4559
Qy	762	sLeuGluGluLysAlaGlu---AspSerSerValGluGlnLeuSerThrGluLeuGluPr	781
Db	4560	TAACGATGCAAAAACAAGAGTTGATATCTCAATTTAAATCAAGCGGAAACAAATATTGACCA	4619
Qy	781	oSerArgGluGluSerTrpVal-----Se	789
Db	4620	ATCTTCAACAGATGAATATGTTGATTAATGCAGTTTAAAGAAGAAAGCTAAATTAATGC	4679
Qy	789	rIleLysLysPheIleProGlyArg-----	798
Db	4680	AGTTAAACATTTAGTGAGTACAAAAGATGCTTTAGCTAAGATTGAGCTGCATATAA	4739
Qy	798	-----	798
Db	4740	TGCTAAAGTAACCGACGGATAACTCTAACGCATCGACTTCAAGTGAATTCGAGAAGC	4799
Qy	799	-----LysLysArgAlaAspGlyLysGlnGluGlnAlaThrVa	811
Db	4800	GAACAACAAAACTTGCTGAATTGAAACAAACCGCGGATCAAAATGTCTCAATCAAGCTACTTC	4859

Qy	811	lgluAspSerGly	proValgluileAsnGlu	aspProAsnVal	-----	826	
Db	4860	TAAAGATGAC	ATTCAAAGTTCAAATTCAT	TAATGACTAGATAATTAACG	ATTACAAT	4919	
Qy	827	-----	-----	-----	-----	840	
Db	4920	TCCAACAGG	TAAANAAGATCAGCT	CAACAGATTATATCGCT	TATGCAGATCAGAGAA	4979	
Qy	840	ggluLysMetGlu	aladlnGlyAsnThrGlu	LeuLeuProGlnLeuGly	AlaValThrVal	860	
Db	4980	AAATAATATT	CAGCTGACACTAATGCA	CAACAGATGAAAGCAAC	CAAGCAATTAAGCA	5039	
Qy	860	lSerGluGluLeu	SerLysThrLeuValHis	ThrValSer	-----	873	
Db	5040	AGTTGACCAA	AAATGTTCAAACCTGC	ATGAGAACTTAATATG	TGTGTGGATAATGTTGA	5099	
Qy	874	-----	-----	-----	-----	885	
Db	5100	CGTTGATG	TCATTAACAGGTTAA	CAAGCAATTCATGCT	TATTCAAGTAGATGCTAC	5159	
Qy	885	rValgluLysArg	SerProSerTrp	IleSerAlaSerValThr	GluProLeuGlu	-----	903
Db	5160	TGTTAAACCT	TAAAGCAACCAAGTG	ATTCATCTAAAGCAGG	AAACCAAGAAATCTAT	5219	
Qy	904	-----	-----	-----	-----	915	
Db	5220	TGATCAAAGT	GACCGATTAACTGCT	GAAGAAAAAATCGA	GCATTAAGCAATGATAA	5279	
Qy	915	uValThrGlu	-----	-----	-----	931	
Db	5280	AATTACAGAT	CAAGCTAAACCAAGG	TATTACTGATGCA	CAACAACTGCTGAAGTTGA	5339	
Qy	931	nThrLeuProGlu	GlyLysAspAlaHis	AspMetValThrSer	GluValAspPhe	Th 951	
Db	5340	AGCGAAAGCT	CAAGACTTGAGC	ATTGATTAACATT	-----	CAAATTGACTCAAC 5390	
Qy	951	rSerGluAlaVal	ThrAlaThrGlu	ThrSerGluAlaLeu	ArgThrGluGluVal	ThrGlu 971	
Db	5391	AGAAAAACA	AAAAAGCTATCGA	GAATTAGAAATTC	CACTAGACCAGATTGA	AGCAGGTGT 5450	
Qy	971	uAlaSerGly	AlaGluGluThr	ThrAsp	-----	980	
Db	5451	AAATGTCG	CGCTGATCTACACT	GAAGAAAAAGACG	TTTACGAATGCTTTAGA	AAGA 5510	
Qy	981	-----	-----	-----	-----	999	
Db	5511	CATTTTAT	CAAGCAACTCA	GATATTTCTGATCA	AACTACAAATGCAG	AAATCGCTAC 5570	
Qy	999	rProValGlnGlu	ValGluSerGlyVal	LeuAspThrGluGluGlu	AluArgGlnThrGlu	1019	
Db	5571	T-----	-----	-----	-----	5618	
Qy	1019	nAlaIle	-----	-----	-----	1035	
Db	5619	AGTAGTTAA	GAAAAAATGCTTTG	GAAGCAATTAGA	GAAGTGTTAACAGCA	AAATAGAAAT 5678	
Qy	1035	1-----	-----	-----	-----	1045	
Db	5679	AATTAANA	ATGCGATGCGATCG	CGGCAAGAAATTC	ACGTACGATGATTTAG	TAGTAG 5738	
Qy	1046	-----	-----	-----	-----	1063	
Db	5739	ATATTTTG	ACAGATTTGCTG	ATAAATTAGATA	AAAAACACAAATAC	CAGAAGTT- 5793	
Qy	1063	aSerGluLysGlu	LysAspValMet	ProLys	-----	GlyProValGlnGluAl 1079	
Db	5794	-GCTGAAT	TACAAAAATGAAC	GATACCTGCA	ATTGAAGCGATTTGT	GCCTCAGAATGATCC 5852	
Qy	1079	aGlyValGlu	HisLeuAlaGlnGly	SerGlu	-----	ThrGlu 1091	
Db	5853	AAATGCTA	ATGATCTAATG	TGGTTTCAGAT	TAATTAATGATGCA	CACGAAATTCMAATGC 5912	
Qy	1091	yGlnAlaThr	ProGluSerLeuGlu	ValProGluValThr	-----	AlaAs 1106	

```

Db 5913 TTAATGCAACACGAGAAATATCTGACAACTTAATGTAATCTGACAAAGATTAATGCCAA 5972
Qy 1106 pValaspHisValAlaThrCys-GlnValIleIysLeu----- 1118
Db 5973 TGCAGATACGCTTCTCAACAACGAAATATATGATGCGGCAACAGTGAAACAAAC 6032
Qy 1119 -----GlnGlnLeuMerGlnGlnAlaValAlaPargIuserSerG 1132
Db 6033 ACCAACATCAGCAAAACGTTGACGCACTGATGATGCAAAATGATAAA-CCGCAAGCTAAATA 6091
Qy 1132 luthrleuThrAspSerGlnThrAsnGlySerThrProLeuAlaAspSerAPThrAla 1152
Db 6092 ACAAT-----TCTAGTGCAGATACATCAACAACAGTCCGCAATGATGATGTAAC 6148
Qy 1152 spGlyThrGlnGlnAspGlnThrIleAsp-----SerGlnAspSerIysAlaThrAlaIle 1171
Db 6149 GCAAAACGAGATGAGATCAATCAATTAATGCGGCAAACTGATAACTGCAACGAAAGCA- 6207
Qy 1171 alArgGlnSerGlnValThrGlnGlnAlaAlaThrAlaGlnIysGlnGlnProSer 1191
Db 6208 -----GATTAATGCAAGCCGCAAGCAAAAGCGCAAAATTAACAATAGTACAACTACG 6259
Qy 1191 hrLeuProAsnAsnValProAlaGlnGlnGlnIleGlnIysGlnGlnProGlyArgAspVal 1211
Db 6260 CAACAAACGAGATGACCAACAGAA----- 6285
Qy 1211 euGlnProThrGlnGlnGlnIleuThrAlaAlaValProValIleuAlaIysThrGln 1231
Db 6286 -----TCAACGCAACGCGCTCCAAACACTGCGAGACA----- 6318
Qy 1231 alGlyGlnIleGlnIleValAspThrLeuAspGlyGlnIysValIysGln-----GlnGln 1250
Db 6319 --GGTGCACAAATCATCTGCTGATGACGAAAGATTAATCATCTGTAATGATTCACAAACAA 6376
Qy 1250 lValPheValHisSerGly-----ProA 1258
Db 6377 ATGCTCAAGTAATTAATAGTCTGCAATCTCAATCAACTAATGCAAGGTTGCAACCA 6436
Qy 1258 snSerGlnIysAlaAspValThrTyrAspSerGlnValMetGlyValAlaGlyCysG 1278
Db 6437 AATCTGAAATTAAGCTAAGGCAAGAAAGATGTGAGATTCACAAATCAAGATAG 6496
Qy 1278 lngIulysGlnSerThrGlnValGlnSerLeuSerGlnGlnGlnGlnMetGlnThr 1298
Db 6497 TTGAATCAACAACCTGAACATGCTCTGACGACATACAGAACCAAAAGATCTTTCTA 6556
Qy 1298 spValGlnIysGlnIysAspGlnThrLysProGlnIleValSerGlnIu-----G 1315
Db 6557 ATACATCAAAAGTAAAGAAAGTAACGACAAAGTCAAAACAGATGCTGAGCAACAAAT 6616
Qy 1315 lylGlnGlnIleuThrAlaAlaProGlnHisGlnIleuThrTyrGly----- 1329
Db 6617 CTGATCAAAACGTTGCATCTAAGCAAGATTAATCTGAAGCAACGTTGATATGATG 6676
Qy 1330 -----LysProValLeuThrLeuAspMetProSerSerGlnIleGlyValIleuG 1347
Db 6677 TATCGAATTAACCA-----TCAACATCTAAACCGTCAGAACCAAAAGATTAAGCG----- 6726
Qy 1347 lylSerIleuGlySerProSerLeuProAspGlnAspIysAlaGlyCysIleGlnValG 1367
Db 6727 -----ACTTCAGTACGAGATGTCACAAAAGGTGACATGCGCAACGCAAG 6769
Qy 1367 lValGlnSerLeuAspThrThrValThrGlnThrAlaGlnAlaValAlaGlyValIleG 1387
Db 6770 ATACAAAGATATCAAGCATCAATGTCGACATGCAAGTGAATTAACATTAAGTACAG 6829
Qy 1387 luthrValIleSerGlnThrGlyGlnSerProGlnCysValGlyAlaHisLeuLeuP 1407
Db 6830 AAAAC-----GAGGTCGCAATGCAATCTC 6853
Qy 1407 roAlGlnIysSerSerAlaThrGlyGlnHisThrThrLeuGlnHisAlaGlnAspThrV 1427

```

```

Db 6854 CAGCT-----ACAG 6862
Qy 1427 alProLeuGlyProGlnSer-----GlnAlaGlnSerIleProIleIleValThrProAlaP 1446
Db 6863 TTTCAAAAGGAGCAATGTGCTTAATCAAGATATGCTTAATGTAATCAAGCTAAAGAA 6922
Qy 1446 roGlnSerThrLeuHisProAspLeuGlnGlnIleSerAlaSerGlnArgIuArgS 1466
Db 6923 ACAAAGCAAAATGCAAAATACAGCCCAACAGAAAGATTAAT-----AAGCTTAAC 6973
Qy 1466 erGlnGlnIleuAspIysProAspAlaGly 1475
Db 6974 AACAAAGTAAAACTTTACCAAGTACTGCT 7002

RESULT 12
US-09-815-242-8291
; Sequence 8291, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyekind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8291
; LENGTH: 7107
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (7107)
US-09-815-242-8291

Alignment Scores:
Pred. No.: 8.8e-20 Length: 7107
Score: 411.00 Matches: 325
Percent Similarity: 35.01% Conservative: 302
Best Local Similarity: 18.15% Mismatches: 671
Query Match: 5.09% Gaps: 495
DB: 10 Gaps: 72

US-09-902-432-4 (1-1596) x US-09-815-242-8291 (1-7107)

Qy 47 ProAlaThrLysLeuProGlnIleuAsnGlyGlnLeuSerSerValAsnGlyValAlaGln 66
Db 2029 CCAAAGTTAACTCAAGCGCAAAACAGATTTATCCAAAGCTTAACAATCTGTAACAA 2088
Qy 67 GlnGlyAspValHisValGlnGlnIleuGlnGlnIleuGlnGlnGlnGlnGlnGlnGlnGln 85

```

Db 2089 CAA-----ATTAAAAAGTCAAATGCATCATTTACAAGATGAAAAAGATGTAGCA 2136
 Qy 86 AspGluaspValGlyGlnArgGluSerGluaspValargGluLysasp----- 101
 Db 2137 AATGATAAATTTGGTAAATTTGAAACAGGCAATTAAGAGTATTGTCAGCAACAACA 2196
 Qy 102 ArgValGluGluMetAlaAAsenSerThrAlaValGluAspileThrLysAspGlyGln 121
 Db 2197 AATGCCAAGTAGAAGCCATTAAACAAAGCGATCAATGATATTAAATCAAACTGCACCT 2256
 Qy 122 -----GluGluThrSerGluile-----Ile 128
 Db 2257 GCTACAACAGCTAAAGCAGCAGCTCTTGAAGAATTTGACCAAGTGTTCACAGCACAAT 2316
 Qy 129 GluGlnIleProAlaSerGluAsnAsnValGluGluMetValGlnProAlaGluSerGln 148
 Db 2317 GATCAAGCACCTTTAAATCTCTGACAACAATGAAGAAGTAGCGGAAGCTATTGAACGT 2376
 Qy 149 AlaAsn-----AspValGlyPheLysLysValPheLysPheValGlyPheLysPheThr 166
 Db 2377 ATTATGCAGCTAAAGTTTCTGGTTAAAGCAATTGAAGCGACAAGCTGCACAAGAT 2436
 Qy 167 ValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThr----- 182
 Db 2437 TTAGAAGAGCTTAAACAGCAAGAAATCTCAAAATTTGAAAATATTACTGACTCTACTCAA 2496
 Qy 183 valLysLysaspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGlu 202
 Db 2497 ACAAAAATGATGCTATATGATGAATTAACAAGCAGCACAGCTAGAAAAAATCAAAAT 2556
 Qy 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSer 222
 Db 2557 GCTACAGTTTCAAATGCAACTAATGAAGTAGCAGAGCTGATGACAGTAGGAAGCA 2616
 Qy 223 ThrGluLys-----GlnGluGlyThrLeuLysGlnGluGlnSerSer 236
 Db 2617 GCTCAAAAGCAAGTTTACATGACATCCAAAGTTGTTAAATCAAAACAGAGAGTTGCTGAT 2676
 Qy 237 ThrGlu-----IleProLeuGlnAlaGluSerAspGln 247
 Db 2677 ACAAAATCAAAAGTATTAGATAAAATCAATGCAATTCAAACACACAGCAAAAGTTAAACCT 2736
 Qy 248 AlaAlaGluGluAlaLysasp----- 255
 Db 2737 GCAGCTGATACGAGAGTAGAAAACGCATATAATACAGTAAACAAGAAATTCAAAATAGC 2796
 Qy 256 -----GluGlyGluGluLysGln-----GluLysGluProThrLysSerPro 269
 Db 2797 AATGCTCAACTACAGAAAGAAAACAAGCTGCATATACAGAAATTAGACTATAAAGCAA 2856
 Qy 270 GluSerProSerSerProValAsnSerGluThrThrSerSerPheLysLysPheThr 289
 Db 2857 GAAGCAAGAACAAATCTTGATGCTGCAAAATACAAACAGTGTATGAACACAGCTAAAGAC 2916
 Qy 290 HisGlyTrpAlaGlyTrpArgLysLysThrSerPheLysLysSerLysGluAsp-----Asp 308
 Db 2917 AATGGTATTGCTGCTCAATTAATCAAGTACAAGCGGCAACAACCTAAGAAATCGGATGCTAAA 2976
 Qy 309 LeuGluThrAlaGluLysArgLysGluGlnGluAla-----GluLysValAspGluGlu 326
 Db 2977 GCGGAATCGCTCAAAAGCAAGTAGAAGCTGAACTGAAATTCGAATTAAGCAATGAATGATCG 3036
 Qy 327 GluLysGluLysThrGluProAlaSerGluGluGlnGluProAla----- 341
 Db 3037 ACTACTGAAGAACAAACAGCTGCAAAAGATAAAGTTGATCAAGCAGTAGTTACTGCAAAC 3096
 Qy 342 GluAspThrAspGlnAlaArgLeuSerAlaAspTyrGluLysValGluLeuProLeuGlu 361
 Db 3097 GCTGATATAGATAATGCTGCGCAAAATACTGATGATGATAATGCAAAAAAATCTAATGAA 3156
 Qy 362 AspGlnValGlyAspLeuGluAlaSerSerGluGluLysCysAlaProLeuAlaThrGlu 381
 Db 3157 GCTACAAATCGCAGCCATTACACCTGATGCAATTTGTTAAA-----CCACAGCGAAACAA 3210

Qy 382 ValPheAspGluLysMetGluAlaHisGlnGluValValAlaGluValHisValSerThr 401
 Db 3211 GCAATTGCTGATAAAGTACAAGCGCAAGAAACAGCAAT----- 3249
 Qy 402 ValGluLysThrGluGluGluGlnGlyGlyGlyGlyGlyGlyGlyGlyValValVal 421
 Db 3250 -----GATGCTAATTAACCGTGCACAACA 3273
 Qy 422 GluGlyThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGluValProGln 441
 Db 3274 GAA-----GAAAAAGCAGCTGCGAAACCAACAAGTTCAAACT 3309
 Qy 442 GluAlaGluProAlaGluLeuMetLysSerArgGluMetCysValSerGlyLysAsp 461
 Db 3310 GAAAAAACACAGCTGATACAGCAAT-----GATGGTGC 3345
 Qy 462 HisThrGlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGluGly 481
 Db 3346 CATACAAAT-----GCAGAGTTGAAGCGCTAAAATGCAGAA----- 3384
 Qy 482 IleValSerGluValGluMetLeu-SerSerGlnGluArgIleLysValGlnGlnSerPr 501
 Db 3385 ---ATTGCTAAATTTGAAGCAATTCAGCGAGCAACAACAATTAAGATTAATCGAAAC-- 3439
 Qy 501 oleuLysLysLeuPheSerSerSerGlyLeuLysLysLeuSerGlyLysGlnLysGln 521
 Db 3440 ----AAGCAATTTGCTACGA-----AAGCGAATGAACGTAAAAACAGCAATTCGCT 3483
 Qy 521 yLysArgGlyGlyGlyGlyAspGluGlu-----ProGlyGlyThrGlnHis1 537
 Db 3484 CAAACGCAAGACATT-ACTGCTGAAGAAATTCGCGGCAATTCGGAATGTAGATAATGC 3542
 Qy 537 ehistrGlu-----SerProGluSerAlaAspGluGlu 548
 Db 3543 TGTGACACAGCAAAATAACAACATTCAAGCTGCTAATAGTCAAAATGATGTAGACCAAGC 3602
 Qy 548 nLys-----GlyGluSerSerAlaSer-----SerProGluGluProGluGluThrTh 564
 Db 3603 GAAAAACAACCTGGTGAAGCTAGTATTGATCAAGTAAACACCAACAGTATTATAAANAAGCAAC 3662
 Qy 564 rCysLeu-----GluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGlu 582
 Db 3663 AGCAGTTACAGACGCGGCAAAATAATATTACAGCTGCACTGATGATGATGTTAGTAGATAC 3722
 Qy 582 yThrThrSerAspGlyGluLysLysArgGlyIleThrProTrpAlaSerPheLysLys 602
 Db 3723 AGCGAAAGATGCAAGTAAAAATTCGATTCAAGAGTACACACACAGCAACAGCGGTTAAATC 3782
 Qy 602 s-----MetValThrProLysLysArgValArgArgPr 613
 Db 3783 AAATGCTAAAAATGATGTTGATCAAGCTGTGACAACTCAAAATCAAGCAATTTGATAATAC 3842
 Qy 613 oSerGluSerAspLysGluGlu----- 620
 Db 3843 AACTGGTGTACAACTGAAGAAAAAATAATGCACAAAAGATTTAGTTTTAAAAGCTAAAGA 3902
 Qy 621 -----GluLeuGluLysValLysSe 627
 Db 3903 AAAAGCGTATCAAGATATCTTAAATGCACAAACAACTACGATGTTACGCAATTAAGA 3962
 Qy 627 rAlaThrLeuSer-----SerThrAspSerThrValSerGluMet---Glu 641
 Db 3963 TCAAGCAGTTGCTGATGTTCAAGGTATTACTGCAGATACAACAATTAAGATGTTGCGAA 4022
 Qy 641 nAspGluValLysThrValGlyGluGluGlnLys----- 652
 Db 4023 AGATGAATTTAGCAACAAAAACAAAGGCAACAAAAAGCACTTTATTGCACAACTGCAGATGC 4082
 Qy 653 -----ProGluGluProLysArgArgValAspThrSerValSerTrp----- 666
 Db 4083 GACTACTGAAGAAAAAAGAAACAAACAAATCAACAGTAGATGTCAGAAATTAACCAAGGTAA 4142

QY 667 -----GluAlaLeuIleCysValGlySerSerValysArgLaaR 680
 Db 4143 TCATAATATTGAAATGACAGCATGATGATGAAACATGCAAAAGATTAATCAAT 4202
 QY 680 GlyAla-----SerSerAspArgGlyGlyProArgThr----- 693
 Db 4203 TCACGCAATTGACCCCAATTCAACATCAACAGTGTAAACGAAATGCAAGCGCAATT 4262
 QY 694 -----LeuGlyGlyAspSerHisArgLaaG 702
 Db 4263 GCTAAGTAAATGCAAAATATAATGAAATTAATTAATTAATGAGACTACTAATA 4322
 QY 702 uGluAlaSerLysAsp-----LysGluAlaGlyThrAspAlaVal 715
 Db 4323 AGAAAAAGGTAAACGATATTGGACCACTGAGTGCATATGAAAGAGTTTAATATATAT 4382
 QY 715 LProAlaSerThr-----GlnGluGlu 722
 Db 4383 TAAATGCAAGACTACTACAGTGAATGTAATGCTAAAGATACAGACATCAAAAGT 4442
 QY 722 nAspGlnAlaGlnGlySerSerSerProGluProAlaGlySerProSerGluGlyGlu 742
 Db 4443 TCACCAACTTCATGCAATCCCTGTAAAGACCAAGCGAGTAAACGCTTATGATCAAGC 4502
 QY 742 yValSerThrTPGluSerPheLysArgLeuValThrProArgLysLysSerLysSer 762
 Db 4503 TCGAGCTGAT---AAGAAAAACAAATAGAACAAACCAACCAATGCTCAACAAACAAAT 4559
 QY 762 sLeuGlnGlyLysAlaGlu---AspSerValGlnGlnLeuSerThrGluLeuGluP 781
 Db 4560 TAAACGATGCAAAACAAAGACTGATCTGATTAATTAATCAAGCAAAACAAATATTGACCA 4619
 QY 781 oSerArgGlnGluSerTrpVal-----Se 789
 Db 4620 ATCTTCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4679
 QY 789 rIleLysLysPheIleProGlyArg----- 798
 Db 4680 AGTTAAACATTAGTAGAGTACAAAGATGCTTAGCTAGCTAGCTAGCTAGCTAGCTAG 4739
 QY 798 ----- 798
 Db 4740 TCGTAAAGTACCGAAGCGGATTAATCTTAACGATGATGATGATGATGATGATGATGATG 4799
 QY 799 -----LysLysArgAlaAspGlyLysGlnGlnGlnAlaThrVal 811
 Db 4800 GAAACAAACAACTGCTGATTAATGAAACAAACGCGGATCAAAATGTCATCAACGCTACTTC 4859
 QY 811 LgluAspSerGlyProValGluLeuGlnGluAspArgProAsnVal----- 826
 Db 4860 TAAAGATGACATTGAAGTTCAAAATTCATTAATGATTAATTAATTAATTAATTAATTA 4919
 QY 827 -----ProAlaValAlaProLeuSerGluTrpAsnAlaValGluThr 840
 Db 4920 TCCACACGCTAAACAAAGATCACTACACAGATTTATATGCTATGACAGATGAGAA 4979
 QY 840 gGluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValLysVal 860
 Db 4980 AATATATATTTCGCTGACCTAATGCAACACAAAGATGAAACACAAACGATTTAAGCA 5039
 QY 860 LserGlnGluLeuSerLysThrLeuValHisThrValSer----- 873
 Db 5040 AGTTGACCAAAATGTTCAAACTGATTAAGAAACATTAATTAATGCTGATATGCTGA 5099
 QY 874 -----ValAlaValLysAspGlyThrArgAlaValThr 885
 Db 5100 CCGTATGATGATTAACACACAGTAAACGACAAATGATGATTAATTAATTAATGCTTAC 5159
 QY 885 yValGlnGluArgSerProSerTrpLysSerAlaSerValThrGluProLeuGlu----- 903
 Db 5160 TGTATAACCTTAAGCGCAACAGATGATGATGATTAAGCAAGAGAAACGAAAGATCTAT 5219
 QY 904 -----HisThrAlaGlyGluAlaMetProProValGluGlu 915

Db 5220 TGATCAAGTACCAAGTTAACTGCTGAAGAAAAAAGTAAAGATTAATGATTAACA 5279
 QY 915 uValThrGlu-----LysAspIleIleAlaGluGluThrProVal-----LeuThrGlu 931
 Db 5280 AATTACAGATCAAGCTAAACAAAGGATTAATGATGCAACAACTGCTGAAGTAA 5339
 QY 931 nThrLeuProGluGlyLysAspAlaHisAspAspMetValHisSerGluValAspPhe 951
 Db 5340 AGCGAAAGCTCAAGGACTTGAGGCACTTGTGATACACTT-----CAAAATGACTCAAC 5390
 QY 951 rSerGlnAlaValThrAlaThrLysSerGluAlaLeuArgThrGluGluValThrGlu 971
 Db 5391 AGAAAAACAAAGAGTATCGAAGATTAAGAACTGCACTTAACAGATTAAGACAGGT 5450
 QY 971 uAlaSerGlyAlaGluGluThrThrAsp----- 980
 Db 5451 AAATGTCAGCGCTGATGCTACAACTGAAGAAAAAGAACGTTTACGAATGCTTTAGAGA 5510
 QY 981 ---MetValSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThr 999
 Db 5511 CATTTATCAAAAGCAAGTGAAGATTTCTGATCAAACTAACAAATGCAAGAAATGCTAC 5570
 QY 999 rProValGlnGluValGluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlu 1019
 Db 5571 T-----GTCAAATAATAGCGCTTGAACAACTTAAGACACATGATTAATCC 5618
 QY 1019 nAlaIle-----LeuGlnAlaValAlaAspLysValLysGlnGluSerGlnVal 1035
 Db 5619 AGTATTTAAGAAATGCTTTGAAAGCAATTAGAGAACTGTTTAAACAGCAAAATGAAT 5678
 QY 1035 L-----ProAlaThrGlnThrValGlnAlaGlyThrGly----- 1045
 Db 5679 AATTAAATGACAGTGCAGATGATCGCGCAAAATTCGACCTACGATTAAGTAGAG 5738
 QY 1046 -----SerLysAlaLeuGluLysValGluGluValGluGluAspSerGluValLeuAl 1063
 Db 5739 AATTTTGAACATTTGCTGATTAATTTGATTAACACAAACAAATACAGAAATGTT----- 5793
 QY 1063 sSerGluLysGlnLysAspValMetProLys-----GlyProValGlnGluAla 1079
 Db 5794 -GCTGAATTAACAAATGTAAGACATCTGCAATGACGATGACGATGATGCTCAGATGATCC 5852
 QY 1079 aGlyAlaGluHisLeuAlaGlnGlySerGlu-----ThrGlu 1091
 Db 5853 AATCTATGATGATTAATTAATGATGATGATTAATGATGATGATGATGATGATGATGATG 5912
 QY 1091 yGlnAlaThrProGluSerLeuGluValProGluValThr-----AlaAs 1106
 Db 5913 TAAATGCAACCAAGAAATATCTGCAACAACTTAATGATTAATGATGATGATGATGATG 5972
 QY 1106 pValAspHisValAlaThrCys-GlnValLysLeu----- 1118
 Db 5973 TCGAGATGCTTTCACACAGCAAAATATACAGATGATGCGGACAGGTAACAAAC 6032
 QY 1119 -----GlnGlnLeuMetGlnGlnAlaValAlaProGluSerSer 1132
 Db 6033 AGCAACATGACGCAACAGTTGACGCAACGATGATGATGATGATGATGATGATGATGATG 6091
 QY 1132 LuThrLeuThrAspSerGluThrAsnGlySerThrProLeuAlaAspSerThrAla 1152
 Db 6092 ACAAT---TCTGTGTCAGATGATCAACAAACAGTCCGACAAATGATATGATGATGAT 6148
 QY 1152 spGlyThrGlnGlnAspGluThrLysLeuAsp---SerGlnAspSerLysAlaThrAlaVal 1171
 Db 6149 GCAACACAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6207
 QY 1171 aLysGlnSerGlnValThrGluGluGluAlaAlaThrAlaGlnLysGluGluProSer 1191
 Db 6208 -----GATATGCAACGCGCAGCAAGAAACGCAACAAATTAACATAGTACAACTACAG 6259
 QY 1191 hLeuProAsnAsnValProAlaGlnGlnGluHisGlyGluGluProGluArgAspVal 1211

```
Db 6260 CAACAAAGAGATGACCAACAGAA----- 6285
Qy 1211 euGluProThrGlnGlnGluLeuThrAlaAlaValProValLeuAlaLysThrGluV 1231
Db 6286 -----TCAACAGCAACGGCTCCAACAACCTGCGAGCACA--- 6318
Qy 1231 alGlyGlnGluGluValAspTrpLeuAspGlyGluLysValLysGlu---GluGlnG 1250
Db 6319 --GGTGCAAGATCATCTGCTGATAGCAAGATAATGCATCTGTAATGATTCACAAACAAA 6376
Qy 1250 luValPheValHisSerGly-----ProA 1258
Db 6377 ATGCTGAAGTAATAATAGTGTCAATCTCAATCAACTAATGCGAAGGTGCGACAAACAA 6436
Qy 1258 snSerGlnLysAlaAlaAspValThrTrpAspSerGluValMetGlyValAlaGlyCysG 1278
Db 6437 AATCTGAATAATAAGCTTAAGGCAGAAAAAGATGTGAGATTTCAACAAATCAAAAGTATGG 6496
Qy 1278 lngLysGluSerThrGluValGlnSerLeuSerLeuGluGluGluMetGluThrA 1298
Db 6497 TTGAATCAACAACCTGAACATTTGCTTCAGCAGACATACAGAACCAAAAGATATCTTCTA 6556
Qy 1298 spValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlu-----G 1315
Db 6557 ATACATCAAAAGATAAGAAAGATGACGACAAGTCAACAGATGCTGAGCAACACAAATT 6616
Qy 1315 lyGluGlnGluThrAlaAlaProGluHisGluGlyThrTyGly----- 1329
Db 6617 CTGATACAAACGTTGTCATCTAAACGACAGATAAATCTGAAGGCAACGTTGATACTGATG 6676
Qy 1330 -----LysProValLeuThrLeuAspMetProSerSerGluArgGlyValAlaLeuG 1347
Db 6677 TATCGAATAAACCA-----TCAACATCTAAACCGTCAGAACGCAAAAGATAAAGCG--- 6726
Qy 1347 lySerLeuGlyGlySerProSerLeuProAspGlnAspLysAlaGlyCysIleGluValG 1367
Db 6727 -----ACTTCGACTGAGGATAGTCAAAAGCTGACATGCGCAACAGCAG 6769
Qy 1367 lnValGlnSerLeuAspThrThrValThrGlnThrAlaGluAlaValGluLysValIleG 1387
Db 6770 ATACAAAAGATAATCAAGCATCAATGGTGCACCTGAGATGTAACAAATAAAGCTACGC 6829
Qy 1387 luThrValValIleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeup 1407
Db 6830 AAAAC-----GACGGTGCCTAATGCATCTC 6853
Qy 1407 roAlaGluLysSerSerAlaThrGlyGlyHisTrpThrLeuGlnHisAlaGluAspThrV 1427
Db 6854 CAGCT-----ACAG 6862
Qy 1427 alProLeuGlyProGluSer---GlnAlaGluSerIleProIleIleValThrProAlaP 1446
Db 6863 TTTCAAAAGGAAGCAATAGTGTCTAATCAAGATATGCTTAATGTAACCTAAGCTAAAGAAA 6922
Qy 1446 roGluSerThrLeuHisProAspLeuGlnGlyGluIleSerAlaSerGluArgGluArgS 1466
Db 6923 ACAAGCAAAATGCAAAATACGCCCAACAGCAAAAGTAAT-----AAGCTTAAAC 6973
Qy 1466 erGluGluGluAspLysProAspAlaGly 1475
Db 6974 AACAGCTAAACCTTTACCAGATACTCGT 7002
```

RESULT 13

```
US-09-960-253-145
; Sequence 145, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.556
```

```
; CURRENT APPLICATION NUMBER: US/09/960,253
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 10300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-145

Alignment Scores:
Pred. No.: 2,21e-19 Length: 10300
Score: 408.00 Matches: 315
Percent Similarity: 38.11% Conservative: 305
Best Local Similarity: 19.36% Mismatches: 670
Query Match: 5.05% Indels: 337
DB: 10 Gaps: 60
```

US-09-902-432-4 (1-1596) x US-09-960-253-145 (1-10300)

```
Qy 26 LeuSerGly-----HisGlyProAlaAlaGluAlaSerGlyAlaAla----- 39
Db 2376 CTCTCAGGTGAAGGAACCTTAGCATGTTAACAGAAATTTAGGGCTCAGGTAAGCAACTGGA 2435
Qy 40 GlyAspPro-AlaAspAlaAspProAlaThrLysLeuProGlnLysAsnGlyGlnLeuSe 59
Db 2436 AATGAACCTTGCAGAGCAGAAAGCAAGACATTTGATTATGAAAGC---CAAACATGC 2492
Qy 59 rSerValAsnGlyValAlaGluGlnGlyAspValHisValGlnGluGluAsnGlnGluG 79
Db 2493 CCATGACACCTGCTCCTCTGAACAG-----ATCCATAGTCTCAGCATAGAACCCAAATC 2546
Qy 79 yGlnGluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgG 99
Db 2547 TAAAGATGTGAATTT-----CAAGTTTACAGATGAATGATGATGTCGAG-- 2595
Qy 99 uLysAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAs 119
Db 2596 -----CTTCAGTTTTCTCAGCAGAGATACCTCTAAGAAAGCCCTGCAAAAGCCA 2642
Qy 119 pGlyGlnGluGluThrSerGluIleIle-----GluGlnIleProAlaSerGluAs 136
Db 2643 GCTGCAAAATAGAAAGTGAAGTCTTGAAGGGGCGAGAACGTTGAAGGCATATCTCAAG 2702
Qy 136 nAsnValGluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLy 156
Db 2703 TAAAGTGAAGAACTCTCCAGCGCTTTTCACAGAGGAAGTAATA----- 2751
Qy 156 sValPheLysPheValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAs 176
Db 2752 -----ACAAAAATCGAT----- 2763
Qy 176 pThrValGlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValG 196
Db 2764 -----CAGCTCTTACTAGAAAAAGAGAGATGTGGAACCCCTCCACAAACCATCGA 2816
Qy 196 yAlaGlyAspHisGlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysG 216
Db 2817 GGAGAGGATCAACAGTGCACAGAAATCAGCTTTAGTATGACTGAGAAAAATGTTTCAGCT 2876
Qy 216 uSerGluLeuLysGlnSerThrGlnLysGlnGluGlyThrLeuLysGlnGluGlnSerSe 236
Db 2877 TAATGAAGAGAAAGTTTCTCTCTGGGTTGAAATTAAGACTCTTTAAAGAA----- 2925
Qy 236 rThrGluIleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluAlaLysAspG 256
Db 2926 -----CAGCTAAATTTATATCCAGAGCTGAGGAAGCAAAAAAGAGCAGCGTGAAGAAGA 2981
Qy 256 uGlyGluGluLysGlnGluProThrLysSerProGluSerProGluSerProProVa 276
Db 2982 TAATGAAGTTTCTCTCTGGCCTTAACAAATATATGATGAGATGAGCCAGCAGCAAAAT 3041
Qy 276 lAsnSerGluThrThrSerSerPheLysPhePheThrHisGlyTrpAlaGlyTrpAr 296
```

```

Db 3042 AAGTAAAGAAAGAA-----CTTCAGCATGAATTGACCTTCTGAA 3080
Qy 296 gLySerThrSerPheLeuSerLeu-----GluAspAs 308
Db 3081 GAAAGAAATGAGCAGAGAAAGAAAGCTCCAGCAGCTTTATTAAAGAAAGAGCT 3140
Qy 308 pLeuGluThrAlaGluLeuArgGlyGluGlnGluAlaGluLeuValAspGluGluGly 328
Db 3141 TCTGCAAAAGCTCAGTAGATTGGAGAGAAATTAAGCACTTGAAAGATCTTAAGAA 3200
Qy 328 sGluLeuSerThrGluProAlaSerGluGluGlnGluProAlaGluAspThrAspGlnAlaArg 348
Db 3201 AAAAAATCCCACTCAGTGAAGCTGAGAGGGGAGAGAGTGAAGAGATTAAGAAACAAAGA 3260
Qy 348 gLeuSer-----AlaAspTyrGluLeuValGluLeuProLeuGluAspGlu 363
Db 3261 ATACTCAGAAATAATGTGTGACTTCTTAAGTCCAGAAATTAAGAAATTTATTAAACAGAC 3320
Qy 363 nValGluAspLeuGluAla-----SerSerGluGluLeuCysAl 376
Db 3321 AATATCTGAGAAAGAAAGTGAAGCTACAGCATATTAAGAAAGATTTGAGAAAG----- 3375
Qy 376 aProLeuAlaThrGluValPheAspGluLeuMetGluAlaHisGlnGluValAlaGlu 396
Db 3376 ---CTGGCAGCTGAA-----GAGCAATTCAGGCTCTGTGTAACAA 3413
Qy 396 uValHisValSerThrValGluLeuSerThrGluGluGlnGluGlyGlyGluAlaGlu 416
Db 3414 GATGATCAGACCTTGCCAGATTAACAAACCA----- 3447
Qy 416 uGlyGlyValValGluGlyThrGlyGluSerLeuProProGluLeuLeuAlaGluPr 436
Db 3448 -----ATGATTTGCTCCAGCAAGATCATGTAACAA 3479
Qy 436 oGlnGluValProGlnGluAlaGluProAlaGluGluLeuMetLeuSerArgGluMetCys 456
Db 3480 CCAGCAATTTATCCAGAG-----TTAATCAACAGTAACAGCA 3518
Qy 456 sValSerGlyGlyAspHisThrGlnLeuThrAsp-----LeuSerProGluGlu 472
Db 3519 TCCAACTGATGGGAGCTCCGTAGCACTTGTAAGAAACAGTGTGATTAAGTCCACTTG 3578
Qy 472 uLysThrLeuProLysHisProGluGlyLysValSerGluValGluMetLeuSerSerGlu 492
Db 3579 TACAGGTAGTAGGAACTGGAACCAAGAACTAAGAAAGAAAGATCTGAGCCCTTGAA 3638
Qy 492 nGluArgGlyLeuValGlnGlySerProLeuLysLeuPheSerSerSerGlyLeuLys 512
Db 3639 AAAAAAGAGCACTTCAAAAGAAAGCTAAGAGAGGCTTAACCTCCCGCAAGCAATTCCT 3698
Qy 512 sLysLeuSerGlyLysGlnGlnLysGlyLysArgGlyGlyGlyLysAspGluGluProGlu 532
Db 3699 TAAAAAGGACACAGAAAGAAAGAAAGACATCTCAGG-----GAGAGGCTTAA 3743
Qy 532 yGluTyrGlnHisLeuHisThrGluSerProGluSerAlaAspGluGlnLysGlyGln 552
Db 3744 GCAACAGAAAGATGATTAATGCTTGCAAGAACATTTGATGACAAAGCAAGCAAGAA 3803
Qy 552 r-----SerAl 554
Db 3804 TGAGAAATTTGAGACCAAGCTTAAGCAACTCCAGATTCAAGTAAGGAATTCATAGACG 3863
Qy 554 aSerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGluAlaPr 574
Db 3864 AAAACTCCCAAGACACAGACAGCAAGATCTGTCTTCCATCCCAAGCTTTAGAAAGAAC 3923
Qy 574 oGlnAspGlyGluAlaGluGlnGlyThr----- 584
Db 3924 TTTATTCAAGGCAAGCAAGACATCACTCAACTGTTTAGAGTCCAACTTGTCGCC 3983
Qy 585 -----SerAspGlyGluLysArgGluGlyLysThrProTrrAlaSe 599

```

```

Db 3984 AGACTGGCCTTCTCATTTCTGAAGATGCAAGTGTCTGCAAGGCGGAATCTGTGCCCA 4043
Qy 599 rPheLysLeuMetValThrProLysLysArgValArgArgProSerGluSerAspLysGlu 619
Db 4044 GATTAAAGCCCAAGCTG-----AAAGA 4064
Qy 619 uGluGluLeuGluLysValLys-----SerAlaThrLeuSerSerThrAspSerThrValSe 638
Db 4065 AATTAAGGCTGAGAAAGTAGAGATTGAATTAAGATTGAATGTTCTCAACACAGTGAAGCTTCA 4124
Qy 638 rGluMetGlnAspGluValLysThrValGlyGluGluGlnLysProGluGluProLysArg 658
Db 4125 TAAAAATCAGAAAGGATTTTCACTTCAAGACAGATTAATAACAGGCTTTAGAAAT 4184
Qy 658 gArgValAspThrSerValSerTrrGluAlaLeuLysValGlySerSerLysLysArg 678
Db 4185 CGAGAGCTTAAGACAGATATCCATGAAGCTGAAGTCCATGCCAGACCCGACGACAGAA 4244
Qy 678 gAlaArgLysAlaSerSerSerSerAspAspGluGlyGlyProArgThrLeuGlyLysAsp 698
Db 4245 ATTG-----GAAAGCAGCCCACTACAAATGCTGCGCTAGA 4280
Qy 698 rHisArgAlaGluGluAlaSerLysAspLysGluAlaGlyThrAspAlaValProAlaSe 718
Db 4281 AATCTTAAGAGATTTGCAACCTTAACCTGATGACTG-----CAAAACTCATTAAG 4331
Qy 718 rThrGlnGluAspGlnAlaGlnGlySer-----SerSerProGluProAlaGlu 735
Db 4332 CAAAAGAGAGAGAGAGCTTACTACTCTTGACAACTTAGTGAGAAAGAGAGCTCT 4391
Qy 735 ySerProSerGluGluGlyGlyValSerThrTrrGluSerPheLysArgLeuValThrPr 755
Db 4392 CACTTAAATTAACAGACAGATTAATGAACAAAGAAATTTAAATTAAGGCTGCAATGAC-- 4449
Qy 755 cArgLysLysSerLysSerLysLeuGluGluLysAlaGluAspSerSerValGluGlnLe 775
Db 4450 -----CAGCTAGAAATTCACAGCCAAAGAGCATGTAGAGATTAAGCACT 4496
Qy 775 uSerThrGluLe-----GluProSerArgGluGluSerTrrPva 788
Db 4497 ACAGGTGGAACCTTGTGAATGAAGCAAAACCAAGAGATGAGAGAAAGT----- 4551
Qy 788 lSerLeuLysPheLeuProGlyArgArgLysArgAlaAspGlyLysGlnGluGlu 808
Db 4552 -----AGAGCAAAAGCAAAATTAACAAAGAAACCTGCAAGC 4586
Qy 808 nAlaThrValGluAspSerGlyProValGluLysGlnGluAspAspProAsnValProAl 828
Db 4587 TGCCCTTATTTCCGAAAA-----GAAAGCACTAAAGAAACAAAGTCTTCAAGA 4637
Qy 828 aValValProLeuSerGlyLysAsnAlaValGluArg----- 840
Db 4638 GGAATTTGCTTTGGCCAGAA-----GGTACCATTAAGAGTCTCAACAGTCTGCGCAGATG 4694
Qy 841 -GluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLeuLeuGlyAlaValTyrVa 860
Db 4695 GAAAAAGCCAAAGTTTGTCTCAAAATTAAGAAAGAAAGATACGGTGGAGAGGTTAGCTCT 4754
Qy 860 lSerGluGluLeuSerLysThrLeuValHisThrValSerValAlaValLysAspGlyTh 880
Db 4755 TCTTCAAGAGAGAGAGCAAACTCATTAACAGAAATGACAGAGTCTTTATTGGAAATCA 4814
Qy 880 rArgAlaValThrSerValGluGluArgSerProSerTrrLysSerAlaSerValThrGlu 900
Db 4815 GAGTCTCAGCAGCTCTCTGGA-----AGTCTAAACTAGCTCAAGGGGCTTACAGCA 4868
Qy 900 uProLeuGluHisTrrAlaGlyGluAlaMetProProValGluGluValThrGluLysAs 920
Db 4869 ACACAAGAGAAAGTTAGTAAGAAATTT-----GAATCTTGAATCTTTC 4913
Qy 920 pLysLeuAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAlaHis 940
Db 4914 TTAAGATTGCAAGAACTACT-----GAGTGGCAAGAGAAACA 4949

```



```

: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO: 3423
: LENGTH: 4840
: TYPE: DNA
: ORGANISM: Homo sapiens
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 U76366
: US-09-880-107-3423

```

Alignment Scores:

```

Pred. No.: 4,21e-19 Length: 4840
Score: 398.50 Matches: 382
Percent Similarity: 34.23% Conservative: 242
Best Local Similarity: 20.95% Mismatches: 700
Query Match: 4.94% Indels: 505
DB: 10 Gaps: 78

```

US-09-902-432-4 (1-1596) x US-09-880-107-3423 (1-4840)

```

Qy 10 ArgSerProGluGlnProAlaGlySerAspThrProSerGluLeuValLeuSerGlyHis 29
Db 42 CGTCGAGGAGCGCGCGCGCGCGGT-----GCGCGGTAT 77
Qy 30 GlyProAlaAlaGluAlaSerGlyAlaAlaGlyAspProAlaAspAlaAspProAlaThr 49
Db 78 GCCCAGGCGCAGAGAGCGCGGAGCTACTTCC-----CCTGATCTACACCATCTGCT 131
Qy 50 -LysLeuProGlnLysAspGlyGlnLeuSerSerValaSerGlyValaAlaGlnGlyAs 69
Db 132 GCGGCT-----GCGTATGCGCTGCGCGCGCGGAGATGAGAGCAGAGCGG 179
Qy 69 P-----ValHisValGlnG 74
Db 180 CCAGAGGTCTTCCTGCTCAGCGCCGTAACTCTTGACATCTATACACCTGGCAACA 239
Qy 74 GlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 94
Db 240 AACCTCAGAGCTTGTGTCGAGACGAGAG--GCAGAGAGATGCGGCATCTGCA----- 291
Qy 94 GlnLysPalaArgLysAspArgValaGlnGlnMetAlaAlaAsnSerThrAlaValaG 114
Db 232 -----GCTAAGAAACCCTGTCAGACCCCATCAGACCTCGAAGCTCGGA 341
Qy 114 uAspIleThrLysAspGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 134
Db 342 AGAG--GAGGAGAAAGCAGAAAGCGAAACCCGAAAGCCACCCCAAGCTGACATCTAC 398
Qy 134 GlnLysAsnVal-----GlnLysMetValGlnProAlaGlnSerG 148
Db 399 CAACCTCAGCTGTCGGGCGCGAGCTTGCATCAAGCATGAAAGAAAGAAAGCAGACAA 458
Qy 148 n-----AlaAsnAspValaGlnPheLysValaPheLysPheVal 161
Db 459 GACAGAGAAAGCTGGCAGACCTGGGAAATTCATGCCACACCTCGCACTGGGAAAGCGST 518
Qy 161 LcLysPheLysPheThrValaLysAspLysAsnGlnLysSerAspThrValaGlnLeu 181
Db 519 GCCCAACCTCTTCTGGGAAAGTCTCCAGAAAGTACAGAGAGCCCTCAGCAAAATCTAC 578
Qy 181 uThrValLysLysAspGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 201
Db 579 GTTGCTCTCAGAACTGAGAGAGAGGCGAGCGTCCCGGCTTTGGAGCTGCT-----GC 632
Qy 201 nGlnProSerValaGlnThrAlaValaGlnLysSerAlaSerLysGlnSerGluLeuLysG 221
Db 633 CAAGCTGGAGATGCTGTCAGCG--GGCCAGGCCACACCTCCAGCAGAGACCTCCAG 689
Qy 221 nSerThrGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 241
Db 690 CTCACAGTATGAGACAGAGCTGAGAGTAAAGGCTCTGAAAAAATCTCTCAGGTC----- 744
Qy 241 uGlnAlaGlnSerAspGlnAlaAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 261

```

```

Db 745 -----AAGCTGCTCAGCCCTTGCAGAGCG-----ACCC 776
Qy 261 nGlnLysGlnProThrLysSerProGlnSerProSerSerProValaAsnSerGlnThr 281
Db 777 TGGGAAAGGGCTACCCACAGACCCCTCGGAGAGCAGAGGCTGTAGCTCCAGACCA 836
Qy 281 rSerSerPheLysLysPheThrHisGlyTrpAlaGlyTrpArgLysLysThrSerP 301
Db 837 GCGAGG----- 843
Qy 301 eLysLysSerLysGlnAspAspLeuGlnThrAlaGlnLysArgLysGlnGlnAlaG 321
Db 844 -----AAGCCAGAGAGACTCAGA 863
Qy 321 uLysValaAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 340
Db 864 GAGCAGCAGCAGAGAGTATCTGAC-----AGTGAAGAGAGACGCCAGC 908
Qy 341 -AlaGlnAspThrAspGlnAlaArgLeuSerAlaAspTyrGlnLysValaGlnLeuPro 360
Db 909 TGCCAAAGCCCTGCTTCAGCGCGAAGGCTCAGAAA----- 945
Qy 360 GlnLysPalaGlnLysAspLeuGlnAlaSerSerGlnGln 373
Db 946 -ACCTTCAAGTCGAGCTGCTGCTTGCAGAGAGTCCCAAGAGAGAGAGAGCTGC 1004
Qy 374 -----LysCysAlaProLeuAlaThrGlnValaPheAspGlnLysMetG 388
Db 1005 CCGAGCGCCCTCGGAGAGAGAGGCTGAGCTGCCAAGGCCAGCGGGGAGAGCGGGA 1064
Qy 388 uAlaHisGlnGlnValaValaAlaGlnValaHisValSerThrValaGlnGlnGlnGln 408
Db 1065 GAGGAGCTCGAGAGCAGACAGCAGAG-----GATCGCAGCTGAGAGA 1106
Qy 408 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 428
Db 1107 GAGGCGCTGCTCAGCGGAAAGCTTCAGGAGAGAGGCCCCCAGAGCGCGCTCGGC 1166
Qy 428 uPro-----ProGlnLysLeuAlaGlnProGlnGlnValaProGlnGlnAlaG 444
Db 1167 CCTTCCCAAGAGTCCCTCCAGAGAGAGGCTGCCACACT---CCTAGAGAAACAG 1223
Qy 444 uProAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 464
Db 1224 GCTTCAGACCCCGCAG-----GTCCAGGTGGGAGAGCAGAGAGA 1262
Qy 464 nLeuThrAspLeuSerProGlnGlnLysThrLeuProLysHisProGlnGlnGlnGln 484
Db 1263 GACTCAAGAGAGCAGCAGAGAGTACAG-----AG 1295
Qy 484 GlnValaGlnMetLeuSerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 501
Db 1296 TGACAGAGAGAGTCCGAGACCATGATGACAGTCAAGTGAAGCCCTTGGGAGAAAGCCC 1355
Qy 501 GLeuLysLysLeuPheSerSerSerSerGlyLeuLysLysLeuSerGlyLysGlnLysG 521
Db 1356 CCAGGTGAACCTGCTTACCATGAGGATGAGGCGCTTG-----GG 1397
Qy 521 yLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 541
Db 1398 GAAA-----GGCGCGGCGCAGTCCACCTGGGAAAGTG-----GG 1433
Qy 541 rProGlnSerAlaAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 558
Db 1434 GCTTCAACCCCTCAGCCAGAGTGGGAGAGTGGAGAGAGACTCAGAGACAGTATGTA 1493
Qy 558 uGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 578
Db 1494 GGAAG--CATCAGACAGAGATGAGAGAGTCCACAGCTGTGCGCCCGCTCAGAGA 1551
Qy 578 uAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 598

```

Db 1552 AGTCTCTGGGGAACATCTCTCC-----AGGCCAAACCACCTCCA----- 1591
Qy 598 aSerPheLysMetValThrProLysLysArgValArgArg-----ProSe 614
Db 1592 -----GTCTGTCAAGGGCCCCCTCAGAAAGCAGGCGCTGTAGCGTC 1635
Qy 614 rGluSerAspLysGluGluLeuLysValLysSerAlaThrLeuSerThrAs 634
Db 1636 CAGGTCAG-GCTGAAGGCCATGGCAACTCGAGAGCAGCAG--GAGTCGTGGA 1691
Qy 634 pSerThrValSerGluMetGlnAspGluValLysThrValGlyGlu----- 649
Db 1692 CAGTCGGCAGTCAGGAGCACCAGCAGCAGCATGCTGACCTCAGGCANAAACAGCTCT 1751
Qy 650 -----GluGlnLysProGluGluProLysArgValAspThrSerValSerTrpG 667
Db 1752 GAAATTCCTCAGACCAAGCGCTGCCCAAGAAACCAATACCACTGCATCTGCCAAGT 1811
Qy 667 uAlaLeuIleCysValGlySerLysLysArgAlaArgLysAlaSerSerAspAs 687
Db 1812 CGCCCTGTGGAGTGGGCACCCAAACCCCGGAAGAGGAACTCGGACTTCTCCAGC 1871
Qy 687 pGluGlyGlyProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAs 707
Db 1872 AGGCTCATCCCGCTGTGCTGGGGCACCAGAGACCCAGAGAGGATTTCTCAAGCAG 1931
Qy 707 pLysGluAlaGlyThrAspAlaValPro-----AlaSerThrGlnGluGlnAspG 724
Db 1932 TGAGGAATCAGATGAGGAGAGAGACAGAGGTCTTGCAGTAACCGTGGCAGACGAAA 1991
Qy 724 nAlaGlnGlySerSerProGluProAlaGlySer-----ProSerGlu 740
Db 1992 GTCTGTGGGAAGGCTCCAGGTGAAGCAGCCTCAGTCCCTGTAAGGGCTCTTGGG 2051
Qy 740 yGluGlyValSerThrTrpGluSerPheLysArgLeuValThrProArgLys----- 757
Db 2052 GCAAGGAGCTCTCCA-----GTACTCCTGGGAAGACGGGGCC 2090
Qy 758 -----LysSerLysSerLysLeuGluGluLysAlaGluAspSerSerValGlu 774
Db 2091 TACAGTCACCCAGGTGAAGCTGAAAGAGGAGAGACTCTGAGAGCAGTGAAGGAATC 2150
Qy 774 nLeuSerThrGluIleGluProSerArgGluGluSerTrpValSerIleLysPheI 794
Db 2151 AGCAGGTGAGGAAGCAGCTCATCTCCAGCACAGGTGAAACCTCAGTAAGAAACCCCA 2210
Qy 794 eProGlyArgArgLysLysArgAlaAspGlyLysGlnGluGlnAlaThrValGluAspSe 814
Db 2211 GGCCAAAGCCACCCAGCTCCGCCAGAGCACCTTCAGCAAAAGGGHCAAT----- 2262
Qy 814 rGlyProValGluIleAsnGluAspAspProAsnValProAlaValValProLeuSerG 834
Db 2263 -----TCAGCCCTCGGAAAAGTTGTCATCGCAGC 2291
Qy 834 uTyrAsnAlaValGluArgGluLysMetGluAlaGlnGlyAsnThrGluLeuProGlnLe 854
Db 2292 TGCTCAAGCCCAAGCAGAGGTCTCCATCCAGGTCAAGCCACCAAGTGAGAAACCCCA-- 2349
Qy 854 uLeuGlyAlaValTyValSerGluGluLeuSerLys-----ThrLeuValHisThrVa 872
Db 2350 -----AACAGTACCCTTCTGGCGAGGGGCCAGCATCTGTGCCATCTGT 2393
Qy 872 lSerValAlaValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSe 892
Db 2394 GGGGAAGGCGTGGCTACAGCAGCTCAGGCCCCAGACAGAGGCCGAGAGGAGCTCAGGGAG 2453
Qy 892 rTrpIleSerAlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro-- 911
Db 2454 CAGTGAGGAGGAGTCACAGAGTGAGGAGGAGGAGGAGCTGGCTCAGGCGAAGCCTTC 2513
Qy 912 -----ProValGluGluValThrGluLysAs 920
Db 2514 AGGGAAGACCCACAGATCAGAGCTGCTTGGCTCTTGCACAGGAGTCCCCAGGAAGG 2573

Qy 920 pIleIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAla-- 939
Db 2574 G-----GTCGCCCAACACCTCTCT---GGGAAGACAGGCGCTTCGGCTGCCAGGCAGG 2624
Qy 940 ---HisAspAspMetValThrSer-----GluValAspPheThrSerGluAlaValTh 956
Db 2625 GAAGCAGGATGACTCAGGAGCAGCAGCAGGAGGAATCAGACAGTGTGGGAGGACCGGC 2684
Qy 956 rAlaThrGluThrSerGluAlaLeuArg----- 965
Db 2685 AGCTGTGACCTCTGCCCGAGGTGATTAAACCCCCCTGATTTTTTGTGACCCCTAATCGTAG 2744
Qy 965 ----- 965
Db 2745 TCAGCTGGCCAGCTGCTACACCCGACAGCCAGGCTGCAAGCACCCCGAGGAAGGC 2804
Qy 966 -----ThrGluGluValThrGluAlaSerGlyAlaGlu---GluThrThrAspMetVa 982
Db 2805 CCAGGCTCGGAGAGCAGCAGGAGCTCTCTCCGAGAGCAGAGTATGAGGACGTGAT 2864
Qy 982 lSerAlaValSerGlnLeuThrAspSerProAspThrThrGluGluAlaThrProValG 1002
Db 2865 CCCCCTACACATGCTTGACTCTCTGCTCATCAGAACCAATGTGTGTCACCATGCCACTGC 2924
Qy 1002 nGluValGlu-----SerGlyValLeuAspThrGluGluGluGlu 1015
Db 2925 CCACCAAGAATAATGCCCCCAAGCCAGCATGTGTGGGGCCAGCAGCAAGGAGTCCAG 2984
Qy 1015 uArg----- 1016
Db 2985 TCGGATATCAGATGGCAAGAACAGAGGAGCACCGACCTCAGGTGTCAAAGAAAGAACCC 3044
Qy 1017 -----GlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGlu-- 1031
Db 3045 AGCTTCCCTCCCACTGACCCAGGCTGCCCTGAAGGTCTCTCGCCAGAAAGCCAGTGAGGC 3104
Qy 1032 -----GluSerGlnVal----- 1035
Db 3105 TCAGCCTCTGTTCAGGAGCACCCAGCCTTCAAGTGGGGTTGACAGTCTGTGGAAACACT 3164
Qy 1036 -----ProAlaThrGlnThrValGlnArgThrGly----- 1045
Db 3165 CCTGCAACAAGTCCCCAGAGCACCTCCGTCCAGGCCAAAGGACCAACAAAGTCAAGAA 3224
Qy 1046 -----SerLysAlaLeuGluLysValGluGluValG 1056
Db 3225 ACCTAAGCTTCTGAGGTCCAGCAGGCCACCAAGGCCCTGAGAGCTCAGATCAGAGTGA 3284
Qy 1056 uGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyProva 1076
Db 3285 GGACAGCAGCAGACAGTCTTTCAGGGAGTGAGGAAGATGGTGAAGGGCCCCCAGGGG-- 3339
Qy 1076 lGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyLysAlaThrProG 1096
Db 3340 -----GCCAAGTCAGCCCAACAGCTGGTCCCCCCCCCTC 3374
Qy 1096 uSerLeuGlu-----ValProGluValThrAlaAspVal-----AspHisValAl 1111
Db 3375 CAGGACAGAGCCCTGTGTGAGGAGACCCGACGAGAGTCCAGCAGGAGTATGTGTGGC 3434
Qy 1111 aThrCysGlnValIleLysLeuGlnGlnLeuMetGluGlnAlaValAlaProGluSerSe 1131
Db 3435 GCCATCCAGTCTCTC---CTCTCAGGTTATATGACCCCTGGAGCTAAACCCAGCCAATTC 3491
Qy 1131 rGluThr-----LeuThrAspSerGluThrAsnGlySerThrProLeuAl 1146
Db 3492 CCAGGCTCAAAAGCCACTCCCAAGAGTAGATTCCAGCCCTCAGTTCTCTACTCTGGC 3551
Qy 1146 aspSerAspThrAlaAspGlyThrGlnGlnAspGluThrIleAspSerGlnAspSerLy 1166
Db 3552 CGCCAAAGATGACCCAGATGGCAAGCAGGAGGCAAG-----CCCCA 3593

Db 2802 TTCACTGAAGATGGACAGATGAATTTACTCTTTATTTCCAGATAGTACTCAAAAGCAGTTA 2861
 QY 82 GluGluValValAspGluAspValGlyGlnArgLysSerGluAspValArgGluLysAsp 101
 Db 2862 GAGGAGGTACTGATGAAGACATAGACGCCCATGGAAAATTCACAATTAGATTTACGCCA 2921
 QY 102 ArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGlyGln 121
 Db 2922 ACTACATCAACTGGTATGTCAGAAAAGTCAACTTTGAGAGATTCTACA----- 2969
 QY 122 GluGluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnVal---GluGlu 140
 Db 2970 -----ACTGAAGAAAAGTTCCACCTATCAAGCACTCGAAGCCAGTTTATGCAACC 3023
 QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
 Db 3024 ATGGAAGGAAGTCTTTGGGTGAAGTAGAAGATGTGGACCTCTCTAAGCCAGTATCTACT 3083
 QY 161 Val-----GlyPheLysPheThrValLysLysAsp 170
 Db 3084 GTTCCCCAAATTTGCACACACTTCAGAGGTGGAAGGATTAGCATTTGTTAGTTAGTAGC 3143
 QY 171 LysAsnGlu-----LysSerAspThrValGlnLeuLeuThrValLys 184
 Db 3144 ACCAAGAGCCTACTACTATGTAGACTCTTCCCATACCACTCTCTCTCTGTAATTTCCC 3203
 QY 185 LysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHis-----GlnGlu 202
 Db 3204 AAGACAGACTGGGAGGTGTAGTACCTTCTGTTCCATCAGAAGATGAAAGTTCTAGGTGAA 3263
 QY 203 ProSerValGluThrAlaValGlyGluSerAlaSerLysGluSer-----Glu 218
 Db 3264 CCTCTCAAGACATACTGTGTCATTGATCAGACTCGCCTTGAAGCGACTATTCTCCAGAA 3323
 QY 219 LeuLysGlnSerThrGluLysGlnGlyThrLeuLysGlnGlu-----Gln 234
 Db 3324 ACTATGAGAAACAATAATCAGAGGGGAACAATCAGGAAGAATTCCTTGGAAAGAA 3383
 QY 235 SerSerThrGluIleProLeuGlnAla----- 243
 Db 3384 CAGACTGCAGAGAAACCACTGCTCTCAGTCTTACAGTTGGACTCCCAAGGAGCA 3443
 QY 244 -----GluSerAspGlnAlaAlaGluGluAlaLysAspGlu 256
 Db 3444 GTAAACCACTGGATGAACAAGAGGGCGATGGATCAGCATATACAGTCTCTGAAGATGAA 3503
 QY 257 -----GlyGluGlu----- 259
 Db 3504 TTGTTGACAGGTTCTGAGAGGGTCCAGTTTATAGAAACAACCTCCAGTTGGAAAAATTTGAT 3563
 QY 260 -----LysGlnGluLysGlu 264
 Db 3564 CACAGTGTCTTATCCACAGGTGTGTAACTGAGCACAAAAGTGAACAGATGAAGTG 3623
 QY 265 ProThrLysSerProGlu-----SerProSerSerProValAsn 277
 Db 3624 GTAAACATAACACCGCATTTGGGCCAAAAGTATCTTTAAGTCCAGGCGCTGAACAAAA 3683
 QY 278 SerGluThr-----ThrSerSerPheLysLysPhePheThr 289
 Db 3684 TATGAACAGAGAGGTAGTAGTACAACAGGATTTACATCATCTTTGAGTCCCTTTTAGTACC 3743
 QY 290 HisGlyTriAlaGlyTriArgLysLysThrSerPheLysLysSerLysGluAsp---Asp 308
 Db 3744 CACATTACCAGCTTATGGAGAAACCACTACTGAGAAACATCCCTAGAGGATTTGAT 3803
 QY 309 LeuGluThrAlaGluLysArgLysGluGlnAlaGluLysValAspGluGluLys 328
 Db 3804 TTAGGCTCAGGATTTATTTGAAAAGCCCAAGCCACAGAACTCATAGAAATTTTCAACAATC 3863
 QY 329 GluLysThrGluProAlaSer----- 335

Db 3864 AAAGTCACAGTTCCAAGTGTATATTACCACGTGCTTCAGTTCCAGTACAGACTTTCACACA 3923
 QY 336 -----GluGluGlnGlu 339
 Db 3924 ACTTCAGCATTTCAAGCATCTTCCGGCATCACTAAGAAAACCACTCTCATCGACAGGAA 3983
 QY 340 ProAlaGluAspThrAspGlnAlaArgLysSerAlaAspTyArgLysValGluLeuPro 359
 Db 3984 CCTGGTGAAGAAACAACCAAGTGCATGGTAATCATTTGGAGAATCAACATCTCTCATGTTCT 4043
 QY 360 -----LeuGluAspGlnValGly-----AspLeuGluAlaSerSer 371
 Db 4044 CCCACTACCTTTGAGATATTGTAGCCAAGGAAACAGAACCGATATTGATAGAGAGTAT 4103
 QY 372 GluGluLysCysAlaProLeuAlaThrGlu-----ValPheAspGlu 385
 Db 4104 TTCACGACTTCAAGTCTCTCTACACAGCAACAAGACCAACCCCACTGTGGAAGACAAA 4163
 QY 386 LysMetGluAlaHisGlnGluValValAla----- 395
 Db 4164 GAGGCCCTTTGGACCTCAGCGCTTCTACGCCACAGCCCCCAGCAAGACAAAAATTTTAC 4223
 QY 396 ---GluValHisValSerThrValGluLysThrGluGluGlnGlyGlyGlyGlu 414
 Db 4224 CCTGACATTAATGTTTATATTATGAGTGCAGAGAAATAAGACAGTGAATGAGTGAT 4283
 QY 415 AlaGluGlyGlyValValValGluGlyThrGlyGluSerLeuProGluLysLeuAla 434
 Db 4284 -----TTGAGTGTAAATGGTTCATCCAATAGATTCAAGATCT---AAA 4322
 QY 435 GluProGlnGluValProGlnGluAlaGluProAlaGluGluLeuMet----- 450
 Db 4323 GAAGATGAACCTTTGTAGTGAAGAAACAGATCCAGTGCATGATCTAATGGCTGAATTTTA 4382
 QY 451 -----LysSer 452
 Db 4383 CCTGAATTCCTGCACATAATTTGAATAGACCTATACCACAGTGAAGAAATAAGAAAGAA 4442
 QY 453 ArgGluMetCysValSerGlyGlyAspHisThr-----GlnLeuThrAsp 467
 Db 4443 GAAGAAGAGTGTGCAAAATGCTACTGATGTGACAAACCAACCCCATCTGTGCAGTACATAAAT 4502
 QY 468 LeuSerProGluGluLysThrLeuProLysHisProGluGly----- 481
 Db 4503 GGAAGACATCTCGTTACCACTGTGCCCAAGACCCCAAGAGCTGCAGAAGCTAGCGCTGGC 4562
 QY 481 ----- 481
 Db 4563 CAGTTTGAAGAGTGTGCACCTTCTCAGAATTTCTCGGACAGCTCTGAAAGTGTATCTCAT 4622
 QY 482 -----IleValSerGluValGluMet-----LeuSer 490
 Db 4623 CCATTTGTAATAGCCAAAACGGAAATTTGCTACTGTGTGCAACCTAATGAATCTACAGAA 4682
 QY 491 SerGlnGluArgIleLysValGlnGlySer----- 500
 Db 4683 ACAACTGAGTCTCTTGAAGTTATCGAAGCTGAGACTTACCTGAAACATCAGAACAT 4742
 QY 501 -----ProLeuLysLysLeuPheSerSer 508
 Db 4743 TTTTCAGTGGTGAGCCTGATGTTTCCCAAGTCCCATTCATGAGGAATTTGAAGT 4802
 QY 509 SerGlyLeuLysLysLeuSerGlyLysLysGlnLysGlyArgGlyGlyGlyGlyAsp 528
 Db 4803 GGAACAGCCAAAAA-----GGGCAGAAATCAGTCACAGAGAGA-----GAT 4844
 QY 529 GluGluProGlyGluTyArgHisIleHisThrGlu----- 540
 Db 4845 ACTGAAGTTGGTTCATCAGGCACATGAACATACACTGAACCTGTATCTCTGTTTCTCGAAGAG 4904
 QY 541 ---SerProGluSerAlaAspGluGlnLysGlyGluSerSerAla----- 554
 Db 4905 TCTTCAGGAGAGATTGCCATTGACCAAGAAATCTCAGAAAAATAGCCTTTGCAAGGGGTACA 4964

QY 555 -----SerSerProGluGluProGluGluThrThrCysLeuGluLysGlyProLeuGlu 572
 Db 4965 GAAGTAACATTGTTGGTGAAGTGAAGAAAAAGTACTTCTGTCACATACATCACTCCACTATA 5024
 QY 573 AlaProGlnAspGlyGluAla-----GluGluGlyThrThrSerAspGlyGlu 588
 Db 5025 GTTCCAAAGTTGTCATCAGCATATGTTTTCAGAGAGAGAAAGCACTTCCCTAATAGA--- 5081
 QY 589 LysLysArgGluGlyIleThrProTIP-----AlaSerPheLysLysMet--- 603
 Db 5082 -----AATCTTGCCAGATGACCTGTTGCTTACCAAGAAAGACTGG 5123
 QY 604 -----ValThrProLysLysArgVal----- 610
 Db 5124 GTAGAAGCAACTCTTAGACAAGTTGTAGAGCTCTCAGGAGTCTTCGATTCACATTTACA 5183
 QY 611 ArgArgProSerGluSerAspLysGluGluLeuGluLysValLysSerAlaThrLeu 630
 Db 5184 GAAGGCTCTGGAGAGACAGAAAGATGAAGATACATGTTCACTGTAAGTATT 5243
 QY 631 SerSerThrAspSerThr-----ValSerGlu 639
 Db 5244 TCACAGAGAAATACTCTGATACACTTACTTACACTAGACAGATGATACAGAA 5303
 QY 640 MetGlnAspGluValLysThrValGlyGluGluLysProGluGluProLysArgArg 659
 Db 5304 ACCTTTTGTAGAGTCTCTGCAACACCATTTATCCAGTTTCTGCAACACCTTCTGCAAA 5363
 QY 660 Val-----AspThrSerValSerThrGluAlaLeuIle 670
 Db 5364 GTGGTGCCCTACCAAGTTTGTAACTGAAGAAACAGACACTTCT---GAGTGGATTTCCAGTACC 5420
 QY 671 CysValGlySerSerLysLysArgAlaArgLysAlaSerSerSerAspGluGlyGly 690
 Db 5421 ACTGTGAGAGAAAGAAAGAAAGAG---GAGGAGGAGGAACT 5459
 QY 691 ProArgThrLeuGlyGlyAspSerHisArgAlaGluGluAlaSerLysAspLysGluAla 710
 Db 5460 ACAGGTACG-----GCT 5471
 QY 711 GlyThrAspAlaValProAlaSerThrGluGluGluAlaSerGln-----AlaGlnGly 727
 Db 5472 TCTCACTTTGAGGTATATTCATCTACACAGATCCGATCAATTAATTTTACCTTGA 5531
 QY 728 SerSerSerProGluProAlaGlySerProSerGluGlyGluGlyValSerThrTrpGlu 747
 Db 5532 TTAGAAAGTCCAAATGTAGTACATCTAGATTCAGGT-----ACCAGGAAA 5579
 QY 748 SerPheLysArgLeuValThrProArgLysLysSerLysSerLysLeuGluGlyLys--- 766
 Db 5580 AGTTTATATGCTCTTACCAACACCAACAGTCTGAAGAGGAAATGACAGATTTACTCTCT 5639
 QY 767 -----AlaGluAspSerValGluGluLeuSerThrGluIleGluProSerArgGlu 784
 Db 5640 GTCTTACAGAAACAATACATTAGAAATTTGGGGCAAG-----ACCACCTGAG 5690
 QY 785 GluSerTrpValSerIleLysLysPheIleProGlyArgArgLysArgAlaAspGly 804
 Db 5691 CACAGACGATC-----CATCAACCTGGG 5714
 QY 805 LysGluGluGlnAlaThrValGluAspSerGlyProValGluLeuAsp----- 822
 Db 5715 GTTCCAGAGAGGCTGACACTCTCCACGTAGTCCCTCTCTTTTATGAGAGAGGCG 5774
 QY 823 -----AspProAsnValProAlaValValProLeuSer 833
 Db 5775 TCTGGAGAGGCTGCTGCCGACCAAGAAACCAACCACTGTTTCTTCATTTTAAACGTA 5834
 QY 834 GluTrpAsnAlaValGluArgGluLys-----MetGlu 844
 Db 5835 GAGTAT---GCAATTCAAGCCGAAAGAAAGATGAGTGGCACTTGTCTCCGCAATGGGAA 5891

QY 845 AlaGlnGlyAsnThrGlu----- 850
 Db 5892 ACTACATTCTCCACTGACCCAAACAGACTGTTTGTAGTACAGTAATGAGACAGTACT 5951
 QY 851 -----LeuProGlnLeuLeuGlyAlaValTrpValSerGluGluLeuSerLys--- 866
 Db 5952 GCTGAAAATATTAACCCAAACATCCAGGAAATGATGATTTACAGAGCCATTGAGAAACCA 6011
 QY 866 ----- 866
 Db 6012 AATTATGGGAGAGAAATAGGGGCTTTTCCACAGGTTTCTTTGAGAGAAATTCAGT 6071
 QY 867 -----ThrLeuValHisThrValSer-----ValAlaVal 876
 Db 6072 GGTGACTTTAGAGAACTACTCAACAGTGTCTATCCCATGCAAAAAGAAACGGTAATG 6131
 QY 877 IleAspGlyThr---ArgAlaValThrSerValGluGluArgSerProSerTrpIleSer 895
 Db 6132 ATGGAAGGCTCTGGAGATGAGATTTAGGAGACCCAGACTTCACTTACAGTACT 6191
 QY 896 AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetPro----- 911
 Db 6192 ACTTCAGTT---CACATGATCATATCTGACTCAGAGAGACCCAGTAGACACCATG 6245
 QY 912 -----ProValGluGluValThrGlu-----Lys 919
 Db 6246 GTACAGCACTTACGCTTCCCTGGGAGAGATTATACATCTCAGCTGAGGCTCAGGTGAG 6305
 QY 920 AspIleAlaGluGluThrProValLeuThrGlnThrLeuProGluGlyLysAspAla 939
 Db 6306 CAACGTGCACAGTCAGACAGCTGCTGTGTT---CCAGTCTTCCAGTGTGCAAGAA 6362
 QY 940 HisAspAspMetValThrSerGluValAsp---PheThrSerGluAlaValThrAlaThr 958
 Db 6363 TTTTCTGTACAGCTTCTCCATTCATTCAGCAAGAGATTGGAGAAAGTGGTACTGTCAAT 6422
 QY 959 Glu-----ThrSerGluAlaLeuArgThrGluGluVal-----ThrGluAlaSer 973
 Db 6423 GAAATGTATAGAAAGTATCCATCTTTTACCAACAGCAAGAAAGTGAAGTCAAACTCCA 6482
 QY 974 GlyAlaGluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAsp 993
 Db 6483 GTAGAGAGAGAGAAATGAAGTACGTGACAGTTCA-----ACAACTTCTCCCA 6536
 QY 994 ThrThrGlu-----GluAlaThrProVal---GlnGluVal 1004
 Db 6537 ACTATAGAGCCAGCCAAATTAATGTCTAGGCAAGAGTCAACCTGTAAACAAGAAAT 6596
 QY 1005 GluSerGlyValLeuAspThrGluGluGluGluArgGlnThrGlnAlaIleLeuGlnAla 1024
 Db 6597 GAAAGTGAACA-----ACATCAGAGAGACAAATTCAGAAAGAAAGTCAATTTAAATCC 6650
 QY 1025 ValAlaAspLysValLysGluGluSerGlnValProAlaThrGlnThrValGlnArgThr 1044
 Db 6651 CCTCAAACTCTCCGCAACAGACAGAAACAACATCTTGATTCACACACTTTACTGAAC 6710
 QY 1045 GlySerLys----- 1047
 Db 6711 GAACCAAAACACAGATTATTCGTACTAACACAAAGAAACCTTAGAGTATGATAAA 6770
 QY 1047 ----- 1047
 Db 6771 GAAATGAGAGAGAAACACTTCTTTAGTTAACATGCTACTCCAGATCCAGATGCAAT 6830
 QY 1048 -----AlaLeuGluLys----- 1051
 Db 6831 GGTGTGAATCTTACACAACCTCCCTGAAGCTACTGAAABTCAATTTTCTTACGT 6890
 QY 1052 -----ValGluGluValGluGluAspSerGluValLeu 1062
 Db 6891 ACTGCATTAGTAAGTAATCTATACAGCTGAACATGTAGTCAACATTCACCAATCAAA 6950
 QY 1063 AlaSerGluLysGluLysAspValMetProLysGly-----ProValGlnGluAla 1079

Db	6951	AAGGAAGAAAGTCAAAACAT--TTTCCGAAGGCATGAGCAACAAATTTCAAGAGTCA	7007
Qy	1080	GlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThr-	1094
Db	7008	GATACTGAGCTCTATTCTCTGAGCTGGGATCAGGAGAAGATTTTAACTACTCTACCA	7067
Qy	1095	ProGluSerLeuGluValProGluVal-	1104
Db	7068	ACAGAGTCAGTGAATTTTACTGAAGTGGAAACAAATCAATAACACATATTATATCCACACAT	7127
Qy	1105	AlaaspValaspHisValAlaThrCysGlnValIleLysLeuGlnGluLeu-	1122
Db	7128	TCTCAAGTGGAAAGTACCTCAAGTCACAAATTTGAAGACATTTTAACACAGAATGGAAATGTG	7187
Qy	1123	GluGlnAlaValAlaProGluSerSerGluThr-	1133
Db	7188	GCAAAAGAGTTGGACCATCTGATCTCAACACAGACATCTTTGAAGGTAGTGGGTCAGTA	7247
Qy	1134	-----LeuThrAspSerGluThrAsnGlySerThr-	1143
Db	7248	ACCAGCAACAATTAATAGAAATTTTAAAGTGACACTGGAGCAGAAGGACCCACCGTGGCA	7307
Qy	1144	ProLeuAlaaspSeraspThrAlaaspGlyThrGlnGlnAspGluThrIleaspSerGln	1163
Db	7308	CCTCTCCCTTTCTCC--ACGGACATCGGACATCTCTCAAAATCAGACTGTCCAGGTGGCA	7364
Qy	1164	AspSerLysAlaThrAlaAlaValArgGlnSerGlnValThrGluGluGlu-	1180
Db	7365	GAAGAAATCCAGACTAGT-----AGACACAACCACTAAGTGAACAGACTCTAACAG	7418
Qy	1181	---AlaAlaThrAlaGlnLysGluGluProSerThrLeuProAsnAsnValProAlaGln	1199
Db	7419	AATTCTTCAACAGCAGAAATTAACGAAACAACAACCTCATCTACTGATTTTCGGCTAGA	7478
Qy	1200	GluGluHisGlyGluGluProGlyArgAspValLeuGluProThrGlnGlnGluLeuThr	1219
Db	7479	---GCTTATGCT-----TTTGAATGGCCAAAGAAATTTGTTACA	7514
Qy	1220	AlaAlaAlaValProValLeuAlaLysThrGluValGlyGlnGlu-	1237
Db	7515	TCAGACCAAAACCATCTGACTGTATTATGAACTTCTGGAGAGAAGATCTGGAGAAGTG	7574
Qy	1238	AspTrpLeuaspGlyGluLysVal-	1249
Db	7575	GATATTGTTGATTCATTTACACTCTCTGCAACTACTCAGGCCAACAGAGAAGCAGC	7634
Qy	1250	GluValPheValHisSerGly-	1264
Db	7635	ACCACATTTGTTCTTGATGGTCCCTCGGAAAAACAATCCTGAGGTGCCAAGCGCTAAAAGCT	7694
Qy	1265	ValThrTrpAsp-----SerGluValMetGlyValAlaGlyCysGlnGlu	1279
Db	7695	GTTACTGCTGATGATTCGCCACAGTTTCAGTGTGCTGCTCTCATTCATTCAGACAGAAC	7754
Qy	1280	LysGluSerThrGluValGlnSerLeu-	1297
Db	7755	AAAAGCTCCCTGATCCAACTACACACTGTCATATACAGTGTCTATAGAGGTCCACA	7814
Qy	1298	AspValGluLysGluLysArgGluThrLysProGluGlnValSerGluGlnGlyGluGln	1317
Db	7815	GACGGTAGTTTCCAAGACCGT-----TTCAGGGAATTCAGGAT	7853
Qy	1318	GluThrAlaAlaProGluHisGluGlyThrTyrGlyLysProValLeuThrLeuaspMet	1337
Db	7854	TCCACCTTAAACCTTAACAGAAAAAACCCACTGAAATAATTTATCATAGACTCCGACAA	7913
Qy	1338	ProSerSerGluArgGlyLysAlaLeuGlySerLeuGlyGlySerProSerLeuProAsp	1357
Db	7914	GAGACAAGGATTTAATTATGACAAATTCACAGAGTACCATCTCTTGAATTTCTACCTGAG	7973
Qy	1358	-----GlnAspLysAlaGlyCysIleGluVal-	1366

Search completed: December 13, 2002, 06:51:00
Job time : 1298 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 13, 2002, 00:22:56 ; Search time 4124 Seconds
(without alignments)
6267.700 Million cell updates/sec

Title: US-09-902-432-4
Perfect score: 8073
Sequence: 1 MGAGSSTQSPQAGSDT.....AWAQRKCLPRLQLKAPVSK 1596

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US0902432/runat.12122002.141452.25812/app_query.fasta_1.1735
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOP=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0902432 @CNG 1 1 3512 @runat.12122002.141452.25812 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_esti:*
- 10: gb_estc:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1268.5	15.7	1269	11	BC015830	BC015830 Homo sapi
2	1041.5	12.9	776	13	BG974884	BG974884 602842771
3	1000	12.4	668	10	B8546856	B8546856 BB546856
4	908	11.2	661	13	B1652823	B1652823 603298753
5	856.5	10.6	674	9	AL045541	AL045541 DXFZP434K
6	854.5	10.6	1117	14	BM804332	BM804332 AGENCOURT
7	853.5	10.6	1153	14	D88456	D88456 D88456 Mouse
8	821	10.2	610	9	AI225639	AI225639 u113009.Y
9	806.5	10.0	646	10	B8618075	B8618075 601462263
10	785.5	9.7	796	9	AI796191	AI796191 wh43905.X
11	782	9.7	873	14	BQ717711	BQ717711 AGENCOURT
12	776.5	9.6	924	14	BQ221404	BQ221404 AGENCOURT
13	748	9.3	530	9	AI156819	AI156819 u144c01.Y
14	744	9.2	679	14	BQ43215	BQ43215 UI-M-EVO-
15	713	8.8	573	12	B400579	B400579 UI-R-CAO-
16	701	8.7	584	12	B448126	B448126 hr8h12.X
17	690.5	8.6	532	10	B350908	B350908 ht63905.X
18	686.5	8.5	886	12	B163549	B163549 601769945
19	686	8.5	641	10	B8665546	B8665546 BB665546
20	682	8.4	875	14	BQ427917	BQ427917 AGENCOURT
21	673.5	8.3	1045	12	B394092	B394092 602456258
22	666.5	8.3	539	12	B835245	B835245 354322 MA
23	657	8.1	644	13	B1446672	B1446672 dac81f11.
24	656	8.1	490	9	AL041664	AL041664 DXFZP434M
25	653	8.1	467	9	AL041405	AL041405 DXFZP434C
26	637.5	7.9	911	12	BG751345	BG751345 602730292
27	632	7.8	923	14	BQ913238	BQ913238 AGENCOURT
28	598.5	7.4	556	12	B3347898	B3347898 dac81f11.
29	592.5	7.3	527	9	AL041657	AL041657 DXFZP434M
30	592	7.3	461	12	B364956	B364956 CM2-NN114
31	591.5	7.3	651	13	BJ044180	BJ044180 BJ044180
32	590.5	7.3	861	13	B1154270	B1154270 602903056
33	590	7.3	459	9	AL040086	AL040086 DXFZP434B
34	577.5	7.2	466	12	B859328	B859328 UI-M-AQO-
35	568.5	7.0	724	14	BM721210	BM721210 UI-E-EQO-
36	566.5	7.0	561	9	AL703803	AL703803 DXFZP686J
37	560	6.9	812	12	BG538601	BG538601 602567330
38	557	6.9	465	12	B364993	B364993 CM2-NN114
39	542	6.7	423	14	W82339	W82339 me98d02.r1
40	539	6.7	446	10	B120002	B120002 UI-R-CAO-
41	537	6.7	431	9	AI322363	AI322363 me98d02.Y
42	513	6.4	1105	14	BQ935865	BQ935865 AGENCOURT
43	506	6.3	883	13	BM047462	BM047462 603628755
44	504	6.2	333	9	AI548287	AI548287 UI-R-C3-t
45	500.5	6.2	436	9	AA159505	AA159505 z072c07.B

ALIGNMENTS

RESULT 1
LOCUS BC015830
DEFINITION Homo sapiens, clone IMAGE:4691862, mRNA.
ACCESSION BC015830
VERSION BC015830.1 GI:16198369
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1269)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2001) National Institutes of Health, Mammalian

QY	1	MetGluValAglSerSerThrGluGluArgSerProGluGln----	ProAlaGlySerAsp	19
Db	25	ATGGAGCCCGGAGACTCCACCGAGCAGCGACCCCGAGACGCCCGCCGAGGGAGCTCC		84
QY	20	ThrProSerGluLeuValLeuSerGlyHisGlyProAlaAlaGluAlaSer--	GlyAla	38
Db	85	ACGCGGGGTGAGCCCGAGCCACGCGCGCGGGCCCTCTGGCCGAGCGCGCGCCAGACACC		144
QY	39	AlaGluAspProAlaAspAla-----AspProAlaThrIleSLeuProGlnIlyAsnGly		56
Db	145	ACCGCGGACCCCGGCATCGCTGCGCTCGGACCCCGCCACCAAGCTCTACAGAAAGAAAGT		204
QY	57	GluLeuSerSerValaenGlyValAlaGluGlnGlyAspValHisValaGlnGluGluAsn		76
Db	205	CAGCTGTCACCATCAATAGGCGGTAGCTGAGCAAGATGAGCTCAGCCTCCAGAGGGTGAC		264
QY	77	GlnGluGlyGln-----	GluGlu	82
Db	265	CTAATATGCCAGAAAGGAGCCCTGACGGTCAAGGACCCCTAAACACCCAGAGGAAGAA		324
QY	83	GluValValaAspGluAspValaGlyGlnArgGluSerGluAspValaArgGluIlyAspArg		102
Db	325	GAAGTCATTGTACAGAGAGTTGGACAGAGAGACTCTGAATGATGGACCAAAAGAGACTCC		384
QY	103	ValGluGluMetAlaAlaAsnSerThrAlaValaGluAspIleThrIlyAspGlyGlnGlu		122
Db	385	GATTAAGAGATGGCTACTAATCAAGCCGGTTTTCACGCATCAACAGATGATGGCCAGAG		444
QY	123	GluThrSerGluIleIleGluGlnIleProAlaSerGluAsnAsnValaGluGluMetVal		142

Db	445	GAGACACCCGAAATTAATCGAAGACATTCTTCACAGAAACCAATTTAGAGAGCTAAC	504
Qy	143	GlnProAlaGluSerGlnAlaAsnAspValGlyPheIysIysValPheIysPheValGly	162
Db	505	CAACCCACTGAGTCCCGAGCTTAATGAAATTTGAATTTGAAGAGGTGTTAAAGTTTGTTGGC	564
Qy	163	PheIysPheThrValIysIysAspIysAsnGluIysSerAspThrValGlnLeuLeuThr	182
Db	565	TTTAAATTCACTGTGGAAGAAAGGATTAAGACAGAGAAAGCTGACATCTGTCCAGTCACTACT	624
Qy	183	ValIysIysAspGluGluGluGluValAlaGluAlaSerValGlyAlaIysAspHisGlnGlu	202
Db	625	GTGAAGAAAGATGAAGGGAGGAGACA-----GCAGGGCTGGCGACCAACAAGAC	675
Qy	203	ProSerValGluThrAlaValAlaGlyIleuSerAlaSerIysGluSerGluLeuIysGlnSer	222
Db	676	CCCAAGCTT-----GGGGCTGGAGAAAGCACATCTCCAAAGAAAGCGAACCCAAACAACT	729
Qy	223	ThrGluIysGlnGluGluIysThrLeuIysGlnGluIysIysSerThrGluAlaProLeuGln	242
Db	730	ACAGAGAAACCCGAAAGAGACCCCTGAAGCTGATGCAAAACCGCAGAAATTTCTCCCCCA	789
Qy	243	AlaGluSerAspGlnAlaAlaGluGluGluAlaIysAspGluGluGluIysGlnGlu	262
Db	790	GCCCAATCTGGCCAAAGCAGTGG--GAGAAATGCAAAAGAGGAGGAGAAAGAAACAAAGAA	846
Qy	263	LysGluProThrIysSerProGluSerProSerSerProValAsnSerGluThrThrSer	282
Db	847	AAAGAACTTACCAAGATCTGCAAGAAATCTCCGACTGCTCCGTAACCAAGTAAACAGAGATCA	906
Qy	283	SerPheIysIysPhePheThrHisGlyTyrAlaGlyTyrAlaGlyIysIysThrSerPheIys	302
Db	907	ACCTTCABAAAATTTCTTCACTCAAGGTTGGGGCCGGCTGGCCCAAAAACCAAGTTTCAGG	966
Qy	303	LysSerIysGluAspAspLeuGluThrAlaGluIysAspIysGluGlnGluAlaGluIys	322
Db	967	AAGCGGAAGGAGATGAAGTGAAGCTTCAAGAGAAAGAAAGAAACAAAGACCCAGAAAAA	1026
Qy	323	ValAspGluGluGluIysGluIysThrGluProAlaSerGluGlu-----	337
Db	1027	GTAAGCACAGAAAGAACGGAAGGCGAAGGCTTCCGCGAAGAACTGACGCGCTCCGAG	1086
Qy	338	-----GlnGluProAlaGluIysAspThrAspArgGlnAlaArgLeuSerAlaAspTyr	353
Db	1087	CAAGCCACCCACAGAGAGCGCGAGAAAGTCCCAACGAGCCCGGTTATCAAGTGAATAT	1146
Qy	354	GluIysValGluLeuProLeuGluAspGlnValGlyAspLeuGluAlaSerSerGluGlu	373
Db	1147	GAGAAAGTTGAGTGGCTCCCTCAGAGAGACCAAGTCAAGTGGCTCCGACAGGGAACCTTCTGAAGAG	1206
Qy	374	LysCysAlaProLeuAlaThrGluValPheAspGluIys	386
Db	1207	AAACCTGCTCCGTTGGCAGCAAGAAAGTGTATTATGAAAAA	1245

RESULT 2
BG974884

LOCUS	BG974884	776 bp	mRNA	linear	EST 12-JUN-2002
DEFINITION	60284277.F1	NCI_CGAP_Mam4	Mus musculus	CDNA clone	IMAGE:4978744 5',

ACCESSION	BG974884
VERSION	BG974884.1
	GI:14362521

SOURCE house mouse.

ORGANISM

REFERENCE 1 (bases 1 to 776)

AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>

TITLE National Institutes of Health, N
FORM Revised (1999)

COMMENT: Contact: Robert Strausberg Ph.D.
Unpublished (1933)

COMPLIANT
CONTACT: NOBEEC BUILDINGS, EMAIL: Ewa11:csaraha-x@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth

Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10975 row: m column: 17
 High quality sequence stop: 753.

FEATURES
source

```

Location/Qualifiers
1..776
/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4978744"
/tissue_lib="NCI CGAP Mam4"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."
```

BASE COUNT 278 a 181 c 216 g 101 t

ORIGIN

Alignment Scores:

Pred. No.:	4,94e-75	Length:	776
Score:	1041.50	Matches:	220
Percent Similarity:	89.27%	Conservative:	13
Best Local Similarity:	84.29%	Mismatches:	22
Query Match:	12.90%	Indels:	6
DB:	13	Gaps:	3

US-09-902-432-4 (1-1596) x BG974884 (1-776)

```

QY 99 GluLyAspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLys 118
DB 1 GAAAGAGACCGAGTAAAGAAATGGCGACCGCAGTCCACAGTGTGAAGATATACAAAG 60

QY 119 AspGlyGlnGluThrSerGluIleGluGlnIleProIleAsnVal 138
DB 61 GACGACGAGGAGAAACACCGGAAATATCGAACAGATCCCTCTTCAGAGCAATGTG 120

QY 139 GluGluMetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysValPhe 158
DB 121 GAAGAATGGCGAGGCTGCTGAGTCCCAAGCTTAAGCTCGGCTTCAAGAGGTATT 180

QY 159 LysPheValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrVal 178
DB 181 AAATTTGTTGTTTAAATTCACGGTGAAGAAGATAAACCAAGTCAGATACCGTC 240

QY 179 GlnLeuLeuThrValLysLysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGly 198
DB 241 CACCTACTACTCTCAAGAGGATGAAGCGGAAGGGGCGAGAGCCCTCCGTCGAGCAGGA 300

QY 199 AspHisGlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGlu 218
DB 301 GACCACAGAGCCGAGGTGGAGACC--GTCGGCGAATCAGATCCCAAGAAAGTGAG 357

QY 219 LeuLysGlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerThrGlu 238
DB 358 CTGAAGCAATCCACAGAGAGCAAGAGGACCCCTGAGCAAGCACAGACACAGAA 417

QY 239 IleProLeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGlyGlu 258
DB 418 ATTCCTCCCTTCAAGCCGAATCTGCTCAAGGAGCCGAGGAAGACGAGCAAGATGGAGAA 477

QY 259 GluLysGlnGluLysGluProThrLysSerProGluSerProSerProValAsnSer 278
DB 478 GAAACCGAGAGAAAGAACCTACCAAGCCCTTAGAATCTCCGACCGCCCTGTGACGAAT 537
```

```

QY 279 GluThrThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLys 298
DB 538 GAGACAAACATCTCTCTTCAAGAAATCTTCTCACTACGGCTGGCGCCGCGCAAGAG 597

QY 299 ThrSerPheLysLysSerLysGluAspAspLeuGluThrAlaGluLysArgLysGluGln 318
DB 598 ACCAGCTTCAAGAAACCAAGGAAGATGATCTGGAACCTTCGAGGAAGAGACGAGGACAA 657

QY 319 GluAlaGluLysValAspGluGluGluLys--GluLysThrGluProAla---SerGlu 336
DB 658 GAGGCTGAAAAAGTAGACCCGAGGAAGCAAGGGGACCAAGACAGAGCCAGCCGCGAG 717

QY 337 GluGlnGlu-ProAlaGlu-AspThrAspGlnAlaArgLeuSer-AlaAspThrGlu 354
DB 718 GAGCAGAGGAGCTGCAGAAAGGGCACAGACGAGCCAGGATGACCCAGCAGCATATGAG 774

RESULT 3
BB546856 688 bp mRNA linear EST 26-OCT-2001
BB546856 RIKEN full-length enriched, 0 day neonate eyeball Mus
musculus cDNA clone E130314119 3' similar to AB020886 Mus musculus
mRNA for SSECKS, mRNA sequence.
BB546856
BB546856.2 GI:16447501
EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 688)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishi,Y., Ito,M., Kawai,J., Konno,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Murakatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
On Jul 31, 2000 this sequence version replaced gi:9618284.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
```


US-09-902-432-4 (1-1596) x BI652823 (1-661)

QY 876 ValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSerThrIleSer 895
 Db 3 GTCAATTGATGGACCAAGCA-GTCAACAGTCCGCAAGAGCGTCCCTTCGTGGATATCT 61
 QY 896 AlaserValThrGluProLeuGluHleThrAlaGlyGluAlaMetProProValGluGlu 915
 Db 62 GCTTCCATGACAGAACTCTTGAGCAGCAGAGGAGTGGCCACACCGCTGTGGAGAG 121
 QY 916 ValThrGluLysAspIleAlaGluGluThrProValLeuThrGlnThrLeuProGlu 935
 Db 122 GTCACGTAAAGACACATCACTCGAGAGCAACTCTGCACTCGCCAGATTTACCAAGG 181
 QY 936 GlyLysAspAlaHisAspMetValThrSerGluValAspPheThrSerGluAlaVal 955
 Db 182 GGCAAGATGCCATGACGACATAGTCACCAAGTCAAGTGGATTTTACCTCAGAAGCAGTG 241
 QY 956 ThrAlaThrGluThrSerGluAlaLeuArgThrGluGluValThrGluAlaSerGlyAla 975
 Db 242 ACAGCCGACAGAAACCAAGAGGCGCTCCGCGCTGAGAACTTACCGAAGCATCAGGGCA 301
 QY 976 GluGluThrThrAspMetValSerAlaValSerGlnLeuThrAspSerProAspThrThr 995
 Db 302 GAAGACACACACATGTTGTCAGATTTCCAGTGTCCAGCTGCCGACCCCGACACCA 361
 QY 996 GluGluAlaThrProValGlnGluValGluSerGlyValLeuAspThrGluGluGlu 1015
 Db 362 GAGGAAGCCACCCAGTTCAGGAGGTAGAGGGTGGCATGTAGTACGGAAGAACAGGAG 421
 QY 1016 ArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysValLysGluGluSerGlnVal 1035
 Db 422 CGCCACAGCGAGCGCGTCTCCAAAGCGTTGCGACAAAGTGAAGAGGAGTCCCGAGTG 481
 QY 1036 ProAlaThrGlnThrValGlnArgThrGlySerLysAlaLeuGluLysValGluGluVal 1055
 Db 482 CTGCAACCCAGACTCTGCGAGAGAGCGGCCGCAAGACATGGAGAGGTGGAGAGGTA 541
 QY 1056 GluGluAspSerGluValLeuAlaSerGluLysGluLysAspValMetProLysGlyPro 1075
 Db 542 GAGGAGACTCCGAGGTGCTGGCTACCGAGAAAGAGAGATGTTGTGCCGGAAGGACCC 601
 QY 1076 ValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGluThrGlyGlnAlaThrPro 1095
 Db 602 GTCAGGAAGCTGAACTGAGCATCTTGCACAGGCTTCCAGAGCTGTACAGGCTACCCCA 661

RESULT 5
 AL045541
 LOCUS AL045541 674 bp mRNA linear EST 29-FEB-2000
 DEFINITION DKF2p434K035_r1 434 (synonym: htes3) Homo sapiens cDNA clone
 DKF2p434K035_5', mRNA sequence.
 ACCESSION AL045541
 VERSION AL045541.1 GI:5433672
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 674)
 Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
 ,S.
 EST (Duesterhoeft, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Duesterhoeft A
 MIPS
 Am Klopferpitz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Olagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No 5' sequence available.
 This clone (DKF2p434K035) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
 1..674
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKF2p434K035"
 /clone_lib="434 (synonym: htes3)"
 /tissue="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
 BASE COUNT 213 a 139 c 212 g 109 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 5,28e-60 Length: 674
 Score: 856.50 Matches: 170
 Percent Similarity: 86.04% Conservative: 21
 Best Local Similarity: 76.58% Mismatches: 30
 Query Match: 10.61% Indels: 1
 DB: 9 Gaps: 1

US-09-902-432-4 (1-1596) x AL045541 (1-674)

QY 424 ThrGlyGluSerLeuProGluLysLeuAlaGluProGlnGluValProGlnGluAla 443
 Db 4 ACAGCAGGGTCTGTGCAGCTGAAGAAATGGTTGAAATGGATGCAGAACCTCAGGAAGCC 63
 QY 444 GluProAlaGluLeuMetLysSerArgGluMetCysValSerGlyGlyAspHisThr 463
 Db 64 GAACCTGCCAAGAGAGTGTGTGAAGCTCAAGAAACGCTGTGTTCGAGAGGACCTTACA 123
 QY 464 GlnLeuThrAspLeuSerProGluGluLysThrLeuProLysHisProGluGlyIleVal 483
 Db 124 CAGGGAGCTGACCTAGTCTGATGAGAAGGTGCTGTCCAAACCCCGAAGCGGTGTG 183
 QY 484 SerGluValGluMetLeuSerSerGlnGluArgIleLysValGlnGlySerProLeuLys 503
 Db 184 AGTGAGTGGAAATGCTGTCATCACAGGAGAGAAATGAAGGTGCAGGAAAGTCCACTAAAG 243
 QY 504 LysLeuPheSerSerSerGlyLysLysLeuSerGlyLysGlnLysGlyLysArg 523
 Db 244 AAGCTTTTACCAGCACTCGCTTAAAGAGCTTCTGGAAGAAACAGAAAGGGAAGA 303
 QY 524 GlyGlyGlyAspGluProGlyGluTyrGlnHisIleHisThrGluSerProGlu 543
 Db 304 ---GGAGGAGGAGACGAGGAATCAGGGGAGCACACTCAGGTTCCAGCCGATTCCTCGGAC 360
 QY 544 SerAlaAspGluGlnLysGlyGluSerSerAlaSerSerProGluGluProGluGluThr 563
 Db 361 AGCCAGGAGGAGCAAAAGGGCGAGAGCTCTGCTCATCCCTCAGGAGCCCGAGGAGATC 420
 QY 564 ThrCysLeuGluLysGlyProLeuGluAlaProGlnAspGlyGluAlaGluGluThr 583
 Db 421 ACGTGTCTCGAAAAGGGCTTAGCCGAGGTGCAGCAGGATGGGGAAGCTGAAGAAGGAGCT 480
 QY 584 ThrSerAspGlyLysLysArgGluGlyIleThrProTrpAlaSerPheLysLysMet 603
 Db 481 ACTTCGATGGAGAGAAAAGAGAGAGTGTCACTCCCTGGGCATCATTCAAAAGAGATG 540
 QY 604 ValThrProLysLysArgValArgArgProSerGluSerAspLysGluGluLeuGlu 623
 Db 541 GTGAGCCCAAGAGCGTGTAGACGGCTTCGAAAGTGATATAAGAGATGAGCTGGAC 600
 QY 624 LysValLysSerAlaThrLeuSerSerThrAspSerThrValSerGluMetGlnAspGlu 643
 Db 601 AAGGTCAAGAGCGCTACCTGTGTCTCCAGGAGAGCACAGCTCTGAAATGCAAGAAGAA 660
 QY 644 ValLys 645
 Db 661 ATGAAA 666

RESULT 6
 BM804332 1117 bp mRNA linear EST 05-MAR-2002
 LOCUS AGENCOURT 6462209 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5581233
 DEFINITION 5', mRNA sequence.
 ACCESSION BM804332
 VERSION BM804332.1 GI:19121155
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1117)
 AUTHORS NIH-MGC http://mgi.mgi.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Place: LLNL12340 row: m column: 10
 High quality sequence stop: 623.
 Location/Qualifiers
 1. 1117
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1b="IMAGE:5581233"
 /clone_1b="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo
 Average insert size 1.75 Kb. Library constructed by Life
 Technologies."
 BASE COUNT 361 a 257 c 339 g 159 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,42e-59 Length: 1117
 Score: 854.50 Matches: 206
 Percent Similarity: 66.58% Conservative: 37
 Best Local Similarity: 56.44% Mismatches: 77
 Query Match: 10.58% Indels: 46
 DB: 14 Gaps: 10
 US-09-902-432-4 (1-1596) x BM804332 (1-1117)
 QY 291 G|TTPALAG|TTPARGLVSLYSTRSERPhelYSLYSerYSGluAAspLeuGlu 310
 DB 1 GGAGCGCTGGCTGGCGCAAAAGACAGTTTCAGGAGCCGAAAGAGATGAGTGGAA 60
 QY 311 ThAAGluLysArgLysGluGluGluAGluLysValAspLugluGluYsGluYs 330
 DB 61 GCTTCAGAGAGAAAGAAAGCAAGAGCCAGAAAGTAAACAGAGAAAGCGAAAG 120
 QY 331 ThrGluProAlaSerGluGlu-----GluGluProAla 341
 DB 121 GCAGAGGTTCCTCCGAGAAATGACCGCTCCGAGCAAGCCACACGAGGCGGCA 180
 QY 342 GluAspThrAspGluAlaArgLysSerAlaAspThrGluLysValGluLeuProLeuGlu 361
 DB 181 GAAAGTCCACAGAGCCCGGTTATGAGTGAATGAGAAAGTGAAGTCCAGAG 240
 QY 362 AspGluValGlyAspLeuGluAlaSerSerGluGluYsCysAlaProLeuAlaThrGlu 381
 DB 241 GAGCAAGTACGTGGCTCCGAGGAGCCTTCGAAAGAGAAACCTGCTCGTGGGACAGAA 300
 QY 382 ValPheAspGluLysMetGluAlaHisGln---GluValAlaAGluValHisValSer 400

DB 301 GTGTTGATGAGAAATATGAAAGTCCACCAAGAGAGTTTGCGCCGAGAGTCCAGCTCAGC 360
 QY 401 ThrValGluYsThrGluGluGluGluGluYsGluYsGluValVal 420
 DB 361 ACCGTGAGAGAGAAAGCAAGAGAGCAG-----AAAAGCGAG 396
 QY 421 ValGluGluYsThrGluGluSerLeuProGluLysLeuAlaGluProGluGluValPro 440
 DB 397 GTGAAAGAAACAGCAGGCTGTGTGCTGACCTGAAAGATTGTGAATGATGACAGAACT 456
 QY 441 GlnGluAlaGluProAlaGluGluLeuMetLysSerArgLysGluMetCysValSerGluYs 460
 DB 457 CAGGAAGCCGAACTGCTCCAGAGAGCTGTGAAGCTCAAGAAAGAGCTGTTCCGAGAG 516
 QY 461 AspHisThrGlnLeuThrAspLeuSerProGluGluYsThrLeuProLysHisProGlu 480
 DB 517 GACCTTACACAGAGAGAGTGAAGTCTGATGAGAAAGTGTGTCACAAACCCCGGAA 576
 QY 481 GlyLeuValSerGluValGluMetLeuSerSerGlnGluAlaGluValGlnGluYs 500
 DB 577 GCGCTTGAGAGAGTGAAGAAAGTGTGATCAGCAGAGAGAAATGAAGTCCAGGGAAGT 636
 QY 501 ProLeuLysLysLeuPheSerSerSerGluYsLysLysSerGluYsLysGluYs 520
 DB 637 CCACCTAAAGAAAGCTTTTACAGCAGCTGCTTAAAGAAAGCTTTCTGGAAAGAAACAGAG 696
 QY 521 GlyLysArgGluYsGluYsAspGluGluProGluGluYsGlnHisIle-HisThrGlu 540
 DB 697 GGAAGAAAG---GAGA-GGAGAGAGAAATCCGGGAGACACACTTCAGTTCCAGGCCGA 752
 QY 540 UserProGluSerAlaAspGluGluYsGluYsSerAlaSerProGluGlu-P 560
 DB 753 TTCTCCGAGCAGCCAGCAAGAGCCAAAGGCGAGAGCTGTGCTCATCCCTCGAGAGAGC 812
 QY 560 roGluGluThrThrLysLysGluYsGluYsProLeuAlaProGluAspGluGluAlaG 580
 DB 813 CGAAGAGAGCCCTGCTGTGAAAGAGGCTTATC-----CAAAG---GGCCACAGAG 863
 QY 580 LngLuglyThrThr-----SerAspGluYsLysLysArgGluYs 594
 DB 864 AATGGGAAACCTGTAAAGAAAGAACTACTTTCAGAGGGGAAAGAAAGAAAGAGGGG 923
 QY 594 IeThrPro---TrpLaserPheLysLys-----MetValThrProLysLysArgVal- 610
 DB 924 CCACCTCCCGGGGGCTTCTTCCAAAGAGAGGGGAGGCCCCCAAAAGAGCGGTGT 983
 QY 611 -----ArgArgProSerGluSerAspLysG 619
 DB 984 TAAAGCGGCCCTTCGGAAGAGGGGTATTAACAAAGAGAACTGGAGACACGGGGTAC 1043
 QY 619 LngLugluLeu 622
 DB 1044 AAGAAAGCCTC 1054
 RESULT 7
 D88456 1153 bp mRNA linear EST 01-APR-1998
 LOCUS D88456 Mouse ICR testis (J.Tsuschida) Mus musculus cDNA clone A-12,
 DEFINITION mRNA sequence.
 ACCESSION D88456
 VERSION D88456.1 GI:1669580
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1153)
 AUTHORS Takeo,M.M., Araki,Y., Matsunaga,A., Yokoi,A., Tsuchida,J., Nishina
 Y., Nozaki,M., Tanaka,H., Koga,M., Uchida,K., Matsunaga, Okuyama
 A., Rochelle,J.M., Nishimune,Y., Matsui,M. and Seldin,M.
 TITLE Mapping of eight testis-specific genes to mouse chromosomes
 JOURNAL Genomics 46 (1), 138-142 (1997)

MEDLINE 98066773
 COMMENT Contact: Junji Tsuchida
 Department of Science for Laboratory Animal Experimentation
 Research Institute for Microbial Diseases, Osaka University
 3-1 Yamadaoka, Suita, Osaka 565, Japan
 Tel: 06-879-8339
 Fax: 06-879-8339
 Email: tsuchida@biken.osaka-u.ac.jp.
 FEATURES Location/Qualifiers
 source 1. .1153
 /organism="Mus musculus"
 /strain="ICR"
 /db_xref="taxon:10090"
 /clone="A-12"
 /clone_lib="Mouse ICR testis (J.Tsuchida)"
 /sex="male"
 /tissue_type="testis" 201 t 1 others
 BASE COUNT 367 a 285 c 299 g 201 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.78e-59 Length: 1153
 Score: 853.50 Matches: 182
 Percent Similarity: 80.08% Conservative: 19
 Best Local Similarity: 72.51% Mismatches: 43
 Query Match: 10.57% Indels: 8
 DB: 14 Gaps: 2
 US-09-902-432-4 (1-1596) x D88456 (1-1153)
 QY 1351 GlySerProSerLeuProAspGlnAspLysAlaGlyCysileGluValGlnValGlnSer 1370
 DB 3 GGAAGCCCTTCTCTCCAGACCAAGACAAAGCAGATTGTCATAGAGGTTCAGTTCAAGC 62
 QY 1371 LeuAspThrValThrGlnThrAlaGluAlaValGluLysValIleGluThrValVal 1390
 DB 63 TCAGACACACCGACTCACTCAACACCGAGCTGTGAAAGAGTCGAGAACTGTGGCA 122
 QY 1391 IleSerGluThrGlyGluSerProGluCysValGlyAlaHisLeuLeuProAlaGluLys 1410
 DB 123 ACTTCAGAGATGGATGAAGATTGTGGAGTGTGCAGGTGCGCAATCATTTACCAGCTGAGAAG 182
 QY 1411 SerSerAlaThrGlyGlyHisThrThrLeuGlnHisAlaGluAspThrValProLeuGly 1430
 DB 183 CTCTCGAAACCGGTGCTACGGAGCTCTTCAGCATGGAGAGACACCGTGCCCGCAGGNG 242
 QY 1431 ProGluSerGlnAlaGluSerIleProIleValThrProAlaProGluSerThrLeu 1450
 DB 243 CCTGAGTCTCAGCAGAGTCTCCATCCCAATATAGTAATCTCTGCTCTGAAAGCATCCTA 302
 QY 1451 HisProAspLeuGlnGlyGluIleSerAlaSerGlnArgGluSerGluGluAsp 1470
 DB 303 CATTTGACCTTCAAGAGAAGTGAGCGCATCCAGAAACAGAGATCAGATGAAGATAAC 362
 QY 1471 LysProAspAlaGlyProAspAlaAspGlyLysGluSerThrAlaIleGluLysValLeu 1490
 DB 363 AAGCCAGATGCTGCTGCTGATGCTGCGGCAAGAGAGTGCAGCAGAGAGAAATCCTC 422
 QY 1491 LysAlaGluProGluIleLeuGluLeuGluSerLysSerHisLysIleValLeuAsnVal 1510
 DB 423 AGGCTGAACCTGAGATCTTGGAACTTGAGATGAAGCAATAGATTGTCCAGAGTGC 482
 QY 1511 IleGlnThrAlaValAspGlnPheAlaArgThrGluThrAlaProGluThrHisAlaTyr 1530
 DB 483 ATCCAGACAGCGTCGACCAAGTTTGACGCTACAGAAACACCCCGGAAACCCAGCTTCT 542
 QY 1531 AspSerGlnThrGlnValProAlaCysArgLeuAspSerArgGluProAsnArgCysTyr 1550
 DB 543 GATTTCAGAAATCAGGTTCCTGTG-----ATGCA-GCCAGGAGGACACACAGATGCTGG 595
 QY 1551 ThrLysMetLys-----AspAlaLysMetLysHisProValProGlnPro 1565
 DB 596 ACAAGATGAAAGCGACCTTCAAGTCTCCCCCAAGATGGAACACATCAGTGCCTAGTCC 655

QY 1566 ArgGluAspLeuGlnValLeuThrValLeuGluAlaThrProAlaGlnProArgLysCysLeu 1585
 DB 656 AGGAGGACTTGGGTTTCTGATAGTTCTGAAGGCATGACGAGGCTTCAGAAATGATCA 715
 QY 1586 ProArgLeuGlnLeuLysAlaProValSerLys 1596
 DB 716 CCACGCTTGCAGTTGAAGTGCAGGTCTCAAG 748
 RESULT 8
 AI225639 610 bp mRNA linear EST 29-OCT-1998
 ujl309.y1 Sugano mouse kidney mkia Mus musculus cDNA clone
 IMAGE:1907968 5' similar to TR:Q62766 Q62766 SSECKS. ;, mRNA
 sequence.
 ACCESSION AI225639
 VERSION AI225639
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 610)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheilenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:976164
 Possible reversed clone: similarity on wrong strand
 Seq primer: custom primer used
 High quality sequence stop: 460.
 Location/Qualifiers
 source 1. .610
 /organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1907968"
 /clone_lib="Sugano mouse kidney mkia"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII (CACATGTG); Site 2: DraIII (CACATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [GTTCGCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGG, 3' site CACATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTAAAGCTGGG and 3' end primer CGACTGCAGCTCGAGACA."
 BASE COUNT 172 a 163 c 185 g 89 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.7e-57 Length: 610
 Score: 821.00 Matches: 167
 Percent Similarity: 88.61% Conservative: 12


```

RESULT 10
AI796191/c
LOCUS      AI796191      796 bp      mRNA      linear      EST 20-DEC-1999
DEFINITION wh43905.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2383544 3',
            similar to TR:000498 000498 MYASTHENIA GRAVIS AUTOANTIGEN GRAVIN ;,
            mRNA sequence.
ACCESSION  AI796191
VERSION     AI796191.1 GI:5361654
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 796)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE       Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabs@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Prepared by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1103 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 458.
FEATURES    Location/Qualifiers
            source          1..796
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:2383544"
                        /clone_lib="NCI-CGAP_Kid11"
                        /lab_host="DH10B"
                        /notes="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
                        a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                        Plasmid DNA from the normalized library NCI CGAP Kid3 was
                        prepared, and as circles were made in vitro. Following HAP
                        purification, this DNA was used as tracer in a subtractive
                        hybridization reaction. The driver was PCR-amplified cDNAs
                        from a pool of 5,000 clones made from the same library
                        (cloneIDs 1322376-1323911, 1456007-1456775, and
                        1500552-1502855). Subtraction by Bento Soares and M.
                        Fatima Bonaldo."
BASE COUNT  140 a 216 c 173 g 266 t 1 others
ORIGIN
Alignment Scores:
Pred. No.:      4,04e-54      Length:      796
Score:          785.50      Matches:    173
Percent Similarity: 70.61%      Conservative: 24
Best Local Similarity: 62.01%      Mismatches: 54
Query Match:    9.73%      Indels:     28
DB:             9      Gaps:      6

US-09-902-432-4 (1-1596) x AI796191 (1-796)

QY 55 AsnGlyGlnLeuSerValAsnGlyValAlaGluGln----- 67
Db 796 AATGGTCTGCTGCNACCATCTATGGCGTAGCTGGACAGATGAGCTCAGCGCTCAGGAGG 737
QY 68 -----GlyAspValHisValGlnGluAsnGlnGluGln 80
Db 736 GTAACCTAGATGTCAGGAAGAGTCCCTGTCGGTCAAGGAGTCTCTAATCAGCAT---GAA 680
QY 81 GluGluGluValValAspGluAspValGlyGlnArgGluSerGluAspValArgGluLys 100
Db 679 GAAGAAGAAAGTCAATTGTCGGCGAGGTTGATCATCATGACTCTGAAGATGTGCGCAGACA 620

```

```

QY 101 AspArgValGluGluMetAlaAlaAsnSerThrAlaValGluAspIleThrLysAspGly 120
Db 619 GACTCCGATACAGAGATGACTACTAAGTCAGCGGTAGTTTCACGACATCATGATGTTGG 560
QY 121 GlnGluGluThrSerGluIleGluGlnIleProAlaSerGluAsnValGluGlu 140
Db 559 CAGGAGAGACACCCGATATATCGAACAGATTCCTTCTTCAGAAAGCAATTAGAGAG 500
QY 141 MetValGlnProAlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPhe 160
Db 499 CTAACAACAACCCACTGAGTCCAGGCTAATGATATTGGATTAAAGAGGTGTTTAAGTTT 440
QY 161 ValGlyPheLysPheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeu 180
Db 439 GTTGGCTTTAAATTCAGTGTGAAAAAGGATAAGACAGAGAGCCTGACACTGTCCAGCTA 380
QY 181 LeuThrValLysLysAspGluGlyGluValAlaGluAlaSerValGlyAlaGlyAspHis 200
Db 379 CTCACGTGTGAGAAAGATGAGGGGAGGAGCA-----GCAGGGGCTGGCGACCAC 329
QY 201 GlnGluProSerValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLys 220
Db 328 CAGGACCCAGCCCTT-----GGGGCTGGAGAAGCAGCATCCAAAGAAAGCGAACCCAAA 275
QY 221 GlnSerThrGluLysGlnGluGlyThrLeuLysGlnGluGlnSerSerThrGluIlePro 240
Db 274 CAATCTACAGAGAAACCCCGAAGAGACCCCTGAAGCGTGAGCAAAAGCCACGAGAAATTTCT 215
QY 241 -LeuGlnAlaGluSerAspGlnAlaAlaGluGluGluAlaLysAspGluGluGluLys 260
Db 214 CCCCCAAGCGGAATCTGGCCACAGCATG---GAGGAATCAAGAGAGAGAGAGAGAA 158
QY 260 sGlnGluLysGluProThrLysSerProGluSerProGluSerProValAsnSerGluTh 280
Db 157 ACAGAAAAAAGAACCTAGCAGTCTCCAGAAATCTCCGACTAGTTCCGTGACCAGTGAAC 98
QY 280 rThrSerSerPheLysLysPhePheThrHisGlyTrpAlaGlyTrpArgLysLysThrSe 300
Db 97 AGGATCAACCTTCAAAAAATTTCTCACTCAAGGTTGGCGCGCTGGCGCAAAAAGGCCAG 38
QY 300 rPheLysLysSerLysGluAspLeuGluThrAlaGluLysArgLysGluGln 318
Db 37 TTTCAGGAAGCCGAGGAG-----GAAAGGAGGAGAGCTTCAG 1

RESULT 11
BO717711
LOCUS      BO717711      873 bp      mRNA      linear      EST 16-JUL-2002
DEFINITION AGENCOURT_8234483 Lupski_sympathetic_trunk Homo sapiens cDNA clone
            IMAGE:6188718 5', mRNA sequence.
ACCESSION  BO717711
VERSION     BO717711.1 GI:21856608
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 873)
            NIH-MGC http://mgs.nci.nih.gov/
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished (1999)
JOURNAL     Contact: Robert Strausberg, Ph.D.
            Email: cgabs@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLNL13584 row: m column: 07
            High quality sequence stop: 627.
            Location/Qualifiers
            source          1..873

```



```

QY 743 ValSerThrTrpGluSerPheLysArgLeuValThrProArgLysLysSerLysSerLys 762
Db 123 GTTTCACCTGGGAGTCAITTAAGAGGTAGTACGCCAGAGAAAAAATCAAAAGTCCAAAG 182

QY 763 LeuGluLysAlaGluAsp-----SerSerValGluGlnLeuSerThrGlu 778
Db 183 CTGGAGAGAAAGCGAGACTCCATAGCTGGGTCTGGTGTAGNACATTCACATCCAGAC 242

QY 779 IleGluProSerArgGluGluSerTrpValSerIleLysLysPheLysProGlyArg 798
Db 243 ACTGAACCCGGTAAGAAGAAATCTGGGTCTCAATCAAGAAGTTTATTCCTGGACGAAG 302

QY 799 LysLysArgAlaAspGlyLysGlnGlnAlaThrValGluAspSerGlyProValGlu 818
Db 303 AGAAAGGCCGATGGGAACCAAGAACAGCCCTGTTGAAGACGCGAGGCCCAACAGGG 362

QY 819 IleAsnGluAspAspProAsnValProAlaValProLeuSerGluTyrAsnAlaVal 838
Db 363 GCACAGCAAGATGACTGTGATGTCCTGGGCTGGTCTCTGCTGATGATGATGCTGTA 422

QY 839 GluArgGluLysMetGlu-----AlaGlnGlyAsnThrGluLeuProGlnLeu 855
Db 423 GAAAGGGAGAAATGGAGGCACAGCAAGCCCAAAAGCGCAGAGCAGCCGACAGCAAG 482

QY 856 GlyAlaValTyrValSerGluLeuSerLysThrLeuValHisThrValSerValala 875
Db 483 GCAGCCACTGAGGTGTCCAGGAGCTCAGCAGAGTCCAGGTTCATATGATGGCAGCAGCT 542

QY 876 ValIleAspGlyThrArgAlaValThrSerValGluGluArgSerProSerTrpIleSer 895
Db 543 GTGCTGACGGGACGAGGAGTACCATTAATTGAAAGAGTCTCTCTCTTGATATCT 602

QY 896 AlaSerValThrGluProLeuGluHisThrAlaGlyGluAlaMetProProValGluGlu 915
Db 603 GCTTCAGTGACAGAACCTCTTGACAAAGTAGAGCTGAAGCCGACACTGTTAACTGAGGAG 662

QY 916 ValThrGluLysPheLysIleAlaGluGluThrPro 927
Db 663 GTATTGAAAGAGAGTAATTGCAGAGAAAGAACCC 698

RESULT 13
AL156819 530 bp mRNA linear EST 30-SEP-1998
LOCUS ui44c01.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
DEFINITION IMAGE:1885248 5' similar to TR:Q62766 Q62766 SSECKS. ;, mRNA
sequence.
ACCESSION AL156819 GI:3685288
VERSION EST.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 530)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:969572
Possible reversed clone: similarity on wrong strand

```

```

Seq primer: custom primer used
High quality sequence stop: 447.
Location/Qualifiers
1. .530
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1885248"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B"
/note="Vector: pME18S-FL3; Site 1: DraIII (CACTGATG);
Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGCGCCCTTTTITTTTTTTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GTGTGGCTACTGG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGATG, 3' site
CACCATGTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTAAAAGCTGG and 3' end primer
CGACCTGCAGCTCGAGACA."
BASE COUNT 151 a 141 c 164 g 74 t
ORIGIN

Alignment Scores:
Pred. No.: 2,84e-51 Length: 530
Score: 748.00 Matches: 152
Percent Similarity: 92.61% Conservatives: 11
Best Local Similarity: 86.36% Mismatches: 13
Query Match: 9.27% Indels: 0
DB: 9 Gaps: 0

US-09-902-432-4 (1-1596) x AT156819 (1-530)

QY 950 PheThrSerGluAlaValThrAlaThrGluThrSerGluAlaLeuArgThrGluGluVal 969
Db 3 TTTTACCTCAGAACAGTGCACAGCGCAGAAACCCACAGAGCGCTCCGCGCTGAAGAACTC 62

QY 970 ThrGluAlaSerGlyAlaGluThrThrAspMetValSerAlaValSerGlnLeuThr 989
Db 63 ACCGAAGCATCAGGGCAGAGAGACACACATGTTGTCTGAGTTTCCCAAGCTGTC 122

QY 990 AspSerProAspThrThrGluGluAlaThrProValGlnGluValGluSerGlyValLeu 1009
Db 123 GACTCCCGGACACACAGAGGAGCCACCCAGTTTCAGAGGTAGAGGGTGGCATGCTA 182

QY 1010 AspThrGluGluGluArgGlnThrGlnAlaIleLeuGlnAlaValAlaAspLysVal 1029
Db 183 GATACGGAAGAACAGGAGCGCCAGAGCGCGCTCTCCAAAGCCGTTGCAGACAAAGTG 242

QY 1030 LysGluLysSerGlnValProAlaThrGlnThrValGlnArgThrGlySerLysAlaLeu 1049
Db 243 AAGAGAGACTCCCAAGTGGCTGCCAAGCCAGACTCTCGAGAGACAGAGCGCCGAAAGCACTG 302

QY 1050 GluLysValGluGluValGluGluAspSerGluValLeuAlaSerGluLysGluLysAsp 1069
Db 303 GAGAGGTGGAGAGGTAGAGGAGGACTCCGAGGTGCTGGCTACCGAAGAGAGAGGAT 362

QY 1070 ValMetProLysGlyProValGlnGluAlaGlyAlaGluHisLeuAlaGlnGlySerGlu 1089
Db 363 GTTGTGCGGAGGAGCCCGCTGCAGGAAGCTGAAACTGAGCATCTTGACACAGGGCTCCGAG 422

QY 1090 ThrGlyGlnAlaThrProGluSerLeuGluValProGluValThrAlaAspValAspHis 1109
Db 423 ACTGTACAGGCTACCCAGAGAGCCTTGAAGTTCTGAAGTTCACAGAGGATGTAGACCGT 482

QY 1110 ValAlaThrCysGlnValIleLysLeuGlnLeuMetGluGlnAla 1125
Db 483 GCCACCATGTCGCCAGGTATTCAAGACCCAGCAGCTGATGGAACAGGCT 530

```


POLYA=Yes.

FEATURES Location/Qualifiers

source 1. 573
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CA0-bhd-a-09-0-UI"
/clone_lib="UI-R-CA0"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pTT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CA0
library is a subtracted library derived from the following
tissues: thalamus, cerebellum, hypothalamus, medulla, pons
, midbrain, cerebral cortex, corpus striatum, testis, and
hippocampus. For a detailed description of the library
from which this clone was derived, please visit our web
site at ratest.eng.uiowa.edu. The subtraction has been
previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 91 a 148 c 124 g 210 t
ORIGIN

Alignment Scores:

Pred. No.:	2,25e-48	Length:	573
Score:	713.00	Matches:	143
Percent Similarity:	84.54%	Conservative:	21
Best Local Similarity:	73.71%	Mismatches:	24
Query Match:	8.83%	Indels:	6
DB:	12	Gaps:	3

US-09-902-432-4 (1-1596) x BF400579 (1-573)

QY	125	SerGluIleIleGluGlnIleProAlaSerGluAsnValGluGluMetValGlnPro	144
DB	564	GCAGAAATATCGACAGATTCCTTCTCAGAAAGCAATTAGAGAGCTAACACACCC	505
QY	145	AlaGluSerGlnAlaAsnAspValGlyPheLysLysValPheLysPheValGlyPheLys	164
DB	504	ACTGAGTCCCGAGGCTAATGATTTGGATTTAAGAAGGTGTTAAGTTTGTGGCTTTAAA	445
QY	165	PheThrValLysLysAspLysAsnGluLysSerAspThrValGlnLeuLeuThrValLys	184
DB	444	TTCACTGTGAAAAGGATAAGACAGACAGCCCTGACACTGTCCAGCTACTCCTGTGAAG	385
QY	185	LysAspGluGlyGluGlyAlaGluAlaSerValGlyAlaGlyAspHisGlnGluProSer	204
DB	384	AAAGATGAAGGGAGGGAGCA-----GCAGGGGCTGGCGACCACAGGACCCCGC	334
QY	205	ValGluThrAlaValGlyGluSerAlaSerLysGluSerGluLeuLysGlnSerThrGlu	224
DB	333	CTT-----GGGGCTGGAGAAGCAGCATCCAAAGAAAGCGAACCCCAACAATCTACAG	280
QY	225	LysGlnGluGlyThrLeuLysGlnGlnSerSerThrGluIleProLeuGlnAlaGlu	244
DB	279	AAACCCGAGAGACCCCTGAGCGCTGAGCAAAAGCCACGCAAAATTTCTCCCCAGCCGAA	220
QY	245	SerAspGlnAlaAlaGluGluAlaLysAspGluGlyGluLysGlnGluLysGlu	264
DB	219	TCTGGCCACAGCAGTG---GAGGAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAA	163
QY	265	ProThrLysSerProGluSerProSerSerProValAsnSerGluThrThrSerSerPhe	284
DB	162	CCTAGCAAGTCTGCAGAAATCTCCGACTAGTCCCGTACCAGTGAACAGGATCAACCTTC	103
QY	285	LysLysPhePheThrHisGlyTyrAlaGlyTyrPargLysLysThrSerPheLysSer	304
DB	102	AAAAAATCTTCACTCAAGGTGGCGCGCTGGCGCAAAAAGACCCAGTTTCAGGAAGCCG	43
QY	305	LysGluAspAspLeuGluThrAlaGluLysArgLysGluGln	318
DB	42	AAGAGGATGAAGTGAAGCTTCAGAGAAAAAAGAAAAA 1	

Search completed: December 13, 2002, 04:10:45
Job time : 4157 secs

THIS PAGE BLANK (USPTO)